

INTELLIGENCE TEST SCORES OF MENTALLY DEFECTIVE PATIENTS AND THEIR RELATIVES

BY L. S. PENROSE

*From the Research Department, Royal Eastern Counties' Institution,
Colchester*

- I. *Introduction* (pp. 1-2).
- II. *The method of determining mental ratio* (pp. 2-9).
- III. *Analysis of Stanford-Binet mental ratios* (pp. 9-15).
- IV. *Analysis of Porteus Maze test results* (pp. 15-17).
- V. *Sex differences* (p. 17).
- VI. *Conclusion* (pp. 17-18).
- References* (p. 18).

I. INTRODUCTION

IN a recently published survey of the family histories of 1280 cases of mental defect⁽¹⁾, the approximate mental capacity of each relative was recorded. Not every relative could be given an intelligence test, and so the ratings of mental capacity which were made in that study had to be based upon a standard obtained from the actual results of testing a large sample. The tests used on this sample population of mentally defective patients and their relatives were the Stanford Revision of the Binet Scale and the Vineland Revision of the Porteus Maze Series. In the present communication, the results of these tests on the sample population are analysed.

The range of chronological age of persons tested was so wide that a special enquiry had to be undertaken to ascertain how the score of a given person was likely to change as he grew older. In the previously published survey only a crude conventional estimate of the intelligence quotient was made. This estimate, though sufficiently accurate to distinguish between normality, mental dullness and mental defect, was not fine enough to enable arithmetical analysis to be undertaken to obtain mean values of the intelligence measurements. The estimation of arithmetic mean values was the intention of the work recorded in the present paper. The method of assigning to each person tested a mental ratio independent of chronological age is described here in § II. A Binet

mental ratio thus calculated was assigned to 682 relatives of 291 patients and a Porteus mental ratio to 481 relatives of 211 patients.¹

The relatives of patients were first ascertained by Miss D. A. Newlyn in the great majority of instances. The testing was carried out by Miss M. V. Matthews and Mr J. C. Raven. Most of the available relatives were children of school age, and thanks are due to the many education authorities who kindly gave facilities for the testing of children in schools. In the preparation of the data for publication, I have had the most valuable assistance of Mr and Mrs J. C. Raven and of Miss H. Lang Brown.

II. THE METHOD OF DETERMINING MENTAL RATIO

In order to be in a position to compare the intelligence of a child with that of an adult, a mental measurement had to be devised which was independent of age. Sometimes the patient was elderly and the relative young; in other instances, the patient was of school age and the relative adult. The Stanford-Binet scale measures mental age with fair accuracy during the years of chronological age from 8 to 12, but it is not standardized for adults. After about the age of 12 or 13, and probably earlier, the average Binet score does not increase at a rate comparable to the increase in chronological age, partly because the individual intelligence ceases to grow so fast and partly because the scale of tests comes to an end. Burt's average figures for performance of children, even on his revised scale, show that normal children of $14\frac{1}{2}$ years of age score only 13 years and 8 months (2). It is sometimes tacitly assumed that the intelligence grows in linear fashion with chronological age up to a point between 14 and 16 years, and that it then suddenly stops growing. Whether or not this is so, mental ability as measured by the Binet I.Q. certainly declines. There are no exact published figures to show where the maximum Binet mental age lies or how it is reached. The writer has therefore made an attempt to measure approximately the growth of ability on the Binet test from early years up till adult life.

Material available for the investigation of the ability to score on the Binet tests at different ages was of two kinds, namely, the single test scores of 644 school children and the records of defective patients who had been retested at intervals of several years. The 644 school children were relatives of patients.² Their mean Binet scores were below the

¹ The detailed results of these tests cannot be included here, but may be obtained on application to Dr L. S. Penrose, Research Department, Royal Eastern Counties' Institution, Colchester.

² There were thirty-eight adult relatives tested.

general average; the mean intelligence quotient tended to decline with age in a manner similar to that observed in data collected by other investigators (Burt⁽²⁾, Herrmann & Hogben⁽³⁾). The age range of the tests was from $4\frac{1}{2}$ to $14\frac{1}{2}$ years (see Table I). The second kind of material consisted of over 1500 pairs of test scores—each pair of scores from one patient in the Royal Eastern Counties' Institution. The interval between the tests varied from 2 to 11 years but was usually 4 or 5 years. The age of the patient at the first test varied from 6 to 31 years. Inspection of

Table I. *Stanford-Binet test scores of 644 school children (relatives of patients)*

Mean chrono- logical age	No. of children	Mean score	Intelligence quotient
5.0	25	4.61	92.2
6.0	55	5.61	93.5
7.0	58	6.40	91.5
8.0	78	7.30	91.2
9.0	77	8.01	89.0
10.0	80	8.90	89.0
11.0	63	9.72	88.4
12.0	65	9.62	80.2
13.0	85	10.95	84.2
14.0	58	11.17	79.8

these figures clearly showed that, although the I.Q. tended to decline as the patient's chronological age increased, the mental age itself actually continued to increase during almost the whole period, though the increase after the age of 20 was slight. The growth of ability on Binet tests was, in this group of defective patients, not unlike the growth of physical stature. It increased rapidly between the ages of 6 and 15, but the rate became gradually less and the maximum was reached after the age of 20. Insufficient data were available to discover in detail what changes took place after the age of 30 years, but it appeared probable that a decline set in which was at first only slight and later became quite marked.

In order to use the material to measure the growth of Binet ability, the assumption was made that the curve of growth for each degree of intelligence was similar to every other; that is to say, the proportional yearly increase (or decrease) was assumed to be independent of grade.¹ It is doubtful whether this assumption is correct where very high or very low grades of intelligence are concerned, but it is probably fairly true for the great majority of intelligence scores recorded in the data. In the present data, the proportional yearly increment in score for a given age

¹ According to Richardson⁽⁴⁾, the S.D. of intelligence is proportional to score plus a constant.

was measured by combining pairs of tests in the following way. For each patient there were two scores, y_1 and y_2 , which had been obtained at ages x_1 and x_2 ¹, respectively. For the age x_1 , the value $\frac{y_2 - y_1}{y_1(x_2 - x_1)}$ was termed 'the subsequent proportional yearly increment'. Similarly, for the age x_2 , $\frac{y_2 - y_1}{y_2(x_2 - x_1)}$ would be 'the previous proportional yearly increment'. For each chronological age x , the proportional increments were collected together and mean values obtained, i.e.

$$\frac{\Sigma (y_2 - y_1)/y}{n} \quad \text{and} \quad \frac{\Sigma (x_2 - x_1)}{n},$$

where n was the number of pairs. At 16 years, for example, there were 159 pairs of tests available for calculating the subsequent proportional yearly increment and 144 pairs for calculating the previous increment: that is, 159 patients of mean age 16 years were retested subsequently and 144 patients of the same mean age had been previously tested. The mean subsequent proportional increment was found to be 0.037 during an average time interval of 5.88 years. The corresponding figures for the previous increment were 0.134 during a period of 3.97 years. From such figures as these, the gradient of the proportional growth at any given age could be estimated.

If s represents the subsequent proportional increment in a period of time x_s and p the previous increase in the period of time x_p , the mean increment per year, at the point from which all the intervals are measured, is

$$\frac{(s \cdot x_p)/x_s + (p \cdot x_s)/x_p}{x_p + x_s}$$

(if the squares of s and p are neglected).² Thus, the gradient of the curve at 16 years (R_{16}) is

$$\left(\frac{0.037 \times 3.97}{5.88} + \frac{0.134 \times 5.88}{3.97} \right) \left(\frac{1}{5.88 + 3.97} \right),$$

and this expression has the value of 0.023.

Similar values were calculated for other points on the scale, namely, at each year of chronological age from 12 to 31.

Between the ages of 6 and 11, it was not possible to obtain enough pairs to estimate accurately the previous increment. The pairs of tests which represented subsequent intervals were, therefore, divided into two

¹ x_1 was always less than x_2 .

² The formula is based on the assumption that the growth curve for any short distance has the properties of a circle.

portions as nearly equal in number as possible, and two estimates of the subsequent increments at different time intervals were made available. For example, at the year 9 there were 34 pairs with intervals 9-11, 9-12 or 9-13 between test and retest, and 42 pairs with intervals of 9-15 or a greater period of years. For the shorter period, whose mean interval was 3.85 years, s had the value 0.260, and, for the longer period (mean interval 7.95 years), the value of s was 0.357. From two readings on the same side of a point the estimation of the gradient at that point is not so accurate as when the readings are on opposite sides of the point; the formula is the same as that for readings taken on either side but the items are subtracted instead of added. Thus, the gradient R_9 has the approximate value

$$\left(\frac{0.260 \times 7.95}{3.85} - \frac{0.357 \times 3.85}{7.95} \right) \left(\frac{1}{7.95 - 3.85} \right),$$

or 0.089.

The values R_6 , R_7 to R_{31} are measures of the gradient on a curve whose relation to the curve of growth of absolute score is that $\frac{(dy/dx)}{y} = R$, where y is the absolute score and x the age of the subject (see Table II). To obtain a curve showing the growth of the absolute score, observed values of R must be integrated and antilogs taken in accordance with the formula

$$\int R_{dx} = \int \frac{1}{y} dy = \log_e yk,$$

where k is a constant. Integration was performed by adding $\frac{1}{2}R_6$, $\frac{1}{2}R_6 + \frac{1}{2}R_7$, $\frac{1}{2}R_7 + \frac{1}{2}R_8$, etc., cumulatively. The natural antilog of the integral values was then obtained. It then only remained to fix, for general purposes, the constant k , which solely determines whether the curve is high or low and which, therefore, is fixed by the average level of mental ability.

From evidence collected by other workers, it appeared that the Binet test gives most reliable and accurate results in the period between 10 and 11 years of chronological age. It was therefore assumed that, for normal individuals, the score at 10.5 years should be fixed at 10.5. After making this adjustment, the curve of growth of intelligence could be drawn. It is admittedly an estimate for normals based on defectives but, since the use to which it is put in the present communication is mainly to deal with people of intelligence somewhat below the normal, this is not a serious fault.

Table II. *Calculation of curve of growth of Binet ability*

Chrono- logical age in years	No. of pairs of tests		$\frac{(dy/dx)}{y} = R$	Cumulant (c) $\Sigma R \Delta x$	$e^c = yk$	y
	Previous interval	Subsequent intervals				
		(i) (ii)				
6.0	—	4 2	+0.12016	0.06008	1.062	6.79
7.0	—	10 7	+0.12828	0.18430	1.202	7.68
8.0	—	23 16	+0.08948	0.29318	1.341	8.57
9.0	—	34 42	+0.08871	0.38228	1.466	9.37
10.0	—	24 28	+0.07559	0.46443	1.591	10.17
11.0	—	37 40	+0.04966	0.52705	1.693	10.82
12.0	16	47 —	+0.05259	0.57818	1.782	11.39
13.0	92	138 —	+0.03815	0.62355	1.865	11.92
14.0	50	66 —	+0.03922	0.66223	1.949	12.45
15.0	77	101 —	+0.02348	0.69358	2.001	12.79
16.0	144	159 —	+0.02267	0.71666	2.046	13.07
17.0	59	113 —	+0.01903	0.73751	2.090	13.35
18.0	41	72 —	+0.01755	0.75580	2.128	13.60
19.0	43	62 —	+0.00908	0.76911	2.158	13.79
20.0	95	78 —	+0.00544	0.77637	2.174	13.89
21.0	148	90 —	+0.00941	0.78380	2.190	13.99
22.0	80	58 —	+0.00559	0.79130	2.207	14.10
23.0	64	60 —	+0.00766	0.79792	2.221	14.19
24.0	72	50 —	+0.00535	0.80443	2.236	14.29
25.0	98	59 —	+0.00261	0.80841	2.245	14.35
26.0	77	50 —	+0.00876	0.81409	2.258	14.43
27.0	69	41 —	+0.00661	0.82178	2.275	14.54
28.0	68	27 —	+0.00399	0.82708	2.287	14.61
29.0	68	26 —	+0.00248	0.83031	2.294	14.66
30.0	56	20 —	-0.00043	0.83133	2.295	14.66
31.0	48	22 —	+0.00446	0.83335	2.301	14.70

To obtain a complete idea of how scores on Binet tests behave over the whole range from 4 to 30 years of chronological age, a composite curve had to be constructed from the two types of data. The tests of the school children reach from 4 to 14 years and the accurate part of the calculated curve reaches from 10 years onward. By assuming, in both cases, that the Binet test gives correct results at 10.5 years, the following comparable scales were obtained: the correspondence between the two scales was fairly good for the years 10–14 where they overlap. The figures, in Table III, for the school children were obtained by multiplying the mean score given in Table I by a factor, 1.125, in order to bring them up to the normal grade at 10.5 years. In order to make use of these figures for the purpose of calculating approximate mental ratios, a smooth curve had to be obtained which would give the supposed true values of normal mean Binet scores at different ages. Between the ages of 5 and 10 for the school children, a hyperbola was fitted. The values obtained from the continuation of this curve below the age of 6 are given in brackets at the top of the right-hand column in Table III. From 10 to 31 years, the

Table III. *Estimates for normal mean scores on Binet test at different chronological ages*

Chronological age in years	Mean scores of school children $\times 1.125$	Calculated values based on pairs of test scores of patients	Composite curve of $\bar{Y}$ (estimated true values)
3	—	—	(3.32)
4	—	—	(4.29)
5	5.18	—	(5.26)
6	6.31	(6.79)	6.25
7	7.20	(7.68)	7.19
8	8.21	(8.57)	8.13
9	9.01	(9.37)	9.09
10	10.01	10.17	10.00
11	10.94	10.82	10.82
12	10.83	11.39	11.53
13	12.32	11.92	12.12
14	12.57	12.45	12.58
15	—	12.79	12.94
16	—	13.07	13.21
17	—	13.35	13.44
18	—	13.60	13.62
19	—	13.79	13.77
20	—	13.89	13.89
21	—	13.99	13.99
22	—	14.10	14.08
23	—	14.19	14.18
24	—	14.29	14.29
25	—	14.35	14.37
26	—	14.43	14.45
27	—	14.54	14.51
28	—	14.61	14.58
29	—	14.66	14.64
30	—	14.66	14.71
31	—	14.70	—
35	—	—	(14.41)
40	—	—	(14.01)
45	—	—	(13.50)
50	—	—	(13.00)
55	—	—	(12.52)
60	—	—	(12.06)

curve was smoothed out by a difference method so that the steps from year to year diminished regularly. Beyond the age of 30, it is probable that Binet scores decline but the degree of decline could not be accurately estimated from the tests and retests of patients. An estimate was, however, necessary for a few elderly patients and figures were chosen which agreed both with the small amount of data derived from tests of older patients and the curves given in a paper by C. C. & W. R. Miles⁽⁵⁾ on the relation of intelligence scores to chronological age from early to late maturity. The continued rise in Binet score, up to 30 years, in the present estimate is comparable to the results obtained by Miles & Miles for tests of verbal experience or arithmetical ability.

For general use in calculating mental ratios (M.R.), the values in the right-hand column of Table III are changed into their corresponding

reciprocals and the inverted values are multiplied by 100. To obtain the M.R. independently of age, the Binet score is multiplied by the appropriate M.R. factor as given in Table IV. The intermediate values of $100/Y$ have to be determined by the method of differences. Thus, suppose a child aged 6 years 9 months to have a Binet score of 5 years 2 months: to obtain the M.R. we find $100/Y$, which is between 14.9 and 13.9, i.e. 14.4 and 14.4 is then multiplied by $5\frac{2}{12}$ or 5.17. The result is an M.R. of 74.

Table IV

Chronological age in years	Mental ratio factor* 100/Y	Chronological age in years	Mental ratio factor* 100/Y
3.0	30.1	16.0	7.57
3.5	26.2	16.5	7.50
4.0	23.3	17.0	7.44
4.5	20.9	17.5	7.39
5.0	19.0	18.0	7.34
5.5	17.4	18.5	7.29
6.0	16.0	19.0	7.26
6.5	14.9	19.5	7.23
7.0	13.9	20.0	7.20
7.5	13.1	21.0	7.15
8.0	12.3	22.0	7.10
8.5	11.6	23.0	7.05
9.0	11.0	24.0	7.00
9.5	10.5	25.0	6.96
10.0	10.0	26.0	6.92
10.5	9.59	27.0	6.89
11.0	9.24	28.0	6.86
11.5	8.94	29.0	6.83
12.0	8.67	30.0	6.80
12.5	8.45	35.0	6.94
13.0	8.25	40.0	7.14
13.5	8.10	45.0	7.41
14.0	7.95	50.0	7.69
14.5	7.83	55.0	8.05
15.0	7.73	60.0	8.29
15.5	7.65		

* To obtain the M.R., multiply the individual's Stanford-Binet score by the factor appropriate to his chronological age.

The Stanford-Binet test gives a good estimate of general intelligence though it gives verbal ability undue weight. For comparison with results obtained by observing the correspondence of Binet ability in patient and relative, a second set of data was obtained in which the subjects were given the Porteus Maze Test (Vineland Revision). An analysis of the growth of Porteus Maze ability was not attempted, but the mean Binet and Porteus scores for 956 individuals, of all ages and grades who were given both tests, were put side by side. The mean score on the Porteus test follows the mean Binet score very closely (Table V) until after the mental age of 10 years; above this mental age the Porteus

score is relatively lower. A small correction is therefore necessary for Porteus scores above 10 years if they are to be treated as equivalent to Binet scores for purposes of calculating mental ratio. The correction actually applied in the present study is given in Table V.

Table V

No. of individuals	Mean Binet score in years	Mean Porteus score in years	Correction factor to be multiplied by Porteus score
10	3.0	3.50	—
23	4.0	3.89	—
44	5.0	5.43	—
100	6.0	6.06	—
142	7.0	7.00	—
152	8.0	8.15	—
146	9.0	9.03	—
127	10.0	10.15	1.00
72	11.0	11.09	1.01
67	12.0	11.56	1.03
43	13.0	12.36	1.05
19	14.0	13.00	1.08
11	15.0	13.41	1.12

III. ANALYSIS OF STANFORD-BINET MENTAL RATIOS

In a previous publication⁽¹⁾ attention has been drawn to the finding that, when propoiti are separated into groups of comparable mental grade, defect is seen to be more common among relatives of simpletons than among relatives of idiots. This inverse regression is clearly shown in the data which are discussed here. Table VI gives the mean M.R. of sibs of patients grouped by similar mental ratio. Patients for whom M.R. ranges from 0 to 29, i.e. idiots and low-grade imbeciles, have sibs whose mean M.R. is about 90. In the lower part of the table the mean M.R. of relatives is seen to be only about 82, and it can be even less than

Table VI. *Stanford-Binet tests : mean mental ratios of sibs compared with mean mental ratios of patients*

No. of pairs	Mental ratio of patient		Mental ratio of relative	
	Range	Mean	Mean	S.D.
21	0-9	5.3	87.5	17.3
31	10-19	16.7	90.0	22.2
28	20-29	24.7	91.3	9.3
21	30-39	34.3	80.1	27.2
19	40-49	44.9	85.4	17.2
31	50-59	54.1	90.1	20.0
28	60-69	63.8	83.6	16.1
33	70-79	72.6	81.5	10.1
6	80-89	85.1	82.2	8.0
3	90-99	91.0	84.3	12.7
221	Total	43.1	86.3	18.1

that of the patients themselves. In Table VII, where relatives other than sibs are included, the same peculiarity is noticeable. One psychopathic patient with a Binet score above the normal average had a child who was feeble-minded: this exceptional case might perhaps have been excluded from the survey.

Table VII. *Stanford-Binet tests: mean mental ratios of relatives (any degree) compared with mean mental ratios of patients*

No. of pairs	Mental ratio of patient		Mental ratio of relative	
	Range	Mean	Mean	S.D.
24	0-9	5.5	88.9	17.0
44	10-19	16.2	91.7	21.1
64	20-29	24.2	94.5	16.7
45	30-39	35.3	83.5	21.9
116	40-49	46.9	84.9	17.4
162	50-59	54.5	86.9	16.3
119	60-69	63.5	87.5	16.9
72	70-79	73.6	81.7	15.4
29	80-89	85.1	85.1	13.8
6	90-99	91.3	82.0	10.1
0	100-109	—	—	—
1	110-119	115.0	65.0	0.0
682	Total	50.2	86.8	17.7

The other phenomenon besides the inverse regression which is characteristic of data of the kind dealt with here is shown by the fact that, though idiots are less likely than mild cases to have defective relatives, they are more likely to have relatives who are idiots. This peculiarity is expressed in Table VI, and especially in Table VII, by the decrease in the standard deviation of the scores of relatives with the rise of the mean score of the patients. Thus, the standard deviation of the scores of relatives of idiots and imbeciles, in Table VII, is seen to be between 17 and 20, whereas the standard deviation of scores of relatives of patients with M.R. 70-99 is between 10 and 15. The nature of the distribution can be easily appreciated by referring to the scatter diagrams (Tables VIII and IX). The distribution of M.R. in relatives of those patients whose own ratios' range of M.R. is from 2 to 37 shows a tendency to bimodality. The scores of the relatives cluster around a horizontal line drawn between 92 and 97, but there are one or two scores far below this figure. The number of very low scores, although small, is greater, in proportion to the number of normals, in this group than in the group of relatives of patients whose own range of M.R. is 39 or over. The distribution of M.R. in the tables can be attributed to a tendency for the relative to resemble the patient accompanied by a tendency for instances of such resemblance to be rarer when the patient is an idiot than when he is a

Table VIII. *Stanford-Binet test; mental ratio of sib compared with mental ratio of patient*

Mental ratio of sib	Mental ratio of patient																			Total
	2	7	12	17	22	27	32	39	42	47	52	57	62	67	72	77	82	87	92	
142	.	.	.	.	.	.	.	.	.	.	1	.	.	.	.	.	.	.	.	1
137	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0
132	.	.	.	.	.	.	.	.	.	.	1	.	.	.	.	.	.	.	.	0
127	.	.	.	.	.	.	.	.	.	.	1	.	.	.	.	.	.	.	.	1
122	.	.	.	1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1
117	.	.	.	3	.	.	.	.	.	1	.	.	.	.	.	.	.	.	.	4
112	.	.	.	2	1	.	1	.	.	.	1	.	.	.	.	.	.	.	.	6
107	.	1	1	4	.	.	.	2	.	.	3	1	.	.	.	.	.	.	.	9
102	2	2	1	.	1	1	1	.	1	2	3	1	.	.	.	.	.	.	.	15
97	.	3	.	2	1	6	1	.	1	.	5	2	5	4	2	3	.	1	1	37
92	2	1	.	3	2	1	2	2	2	2	2	3	.	3	4	1	.	.	1	31
87	1	2	.	3	5	3	1	4	2	.	2	.	3	3	2	1	.	1	.	33
82	1	1	.	.	.	2	2	1	1	1	1	1	4	.	3	1	.	2	.	18
77	.	3	1	4	.	4	2	1	1	2	.	.	1	1	5	.	1	.	.	25
72	1	.	1	2	.	.	.	1	.	1	.	.	.	.	8	1	.	.	.	15
67	.	.	.	1	.	.	.	.	.	.	1	3	.	.	1	.	1	.	1	6
62	.	.	1	1	.	.	.	.	1	1	1	1	1	.	1	1	.	.	.	5
57	.	.	.	1	.	.	.	.	.	1	.	.	1	.	.	1	.	.	.	4
52	.	.	.	.	.	.	.	.	.	.	1	.	1	.	.	.	.	.	.	1
47	.	.	.	.	.	.	.	.	.	1	.	.	1	.	.	.	.	.	.	3
42	.	.	.	.	.	.	.	.	.	1	.	.	1	.	.	.	.	.	.	0
37	.	.	.	.	.	.	.	.	.	1	.	.	.	.	.	.	.	.	.	1
32	.	.	.	.	.	.	.	.	.	.	.	.	.	1	.	.	.	.	.	1
27	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0
22	.	1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1
17	.	.	.	1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1
12	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1
7	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0
Total	7	14	4	27	10	18	9	12	8	11	18	13	16	12	25	8	1	5	3	221

Table IX. *Stanford-Binet test: mental ratio of relative (any degree)*
compared with mental ratio of patient

Mental ratio of relative	Mental ratio of patient																	Total							
	2	7	12	17	22	27	32	37	42	47	52	57	62	67	72	77	82		87	92	97	102	107	112	117
142	.	.	.	.	1	.	.	.	.	.	1	.	.	.	.	.	.	.	.	.	.	.	.	.	2
137	.	.	.	.	.	.	.	.	.	.	.	.	1	.	.	.	.	.	.	.	.	.	.	.	0
132	.	.	.	.	1	.	.	.	.	.	1	.	.	.	.	.	.	.	.	.	.	.	.	.	2
127	.	.	.	.	.	.	.	.	.	.	1	.	.	1	.	.	.	.	.	.	.	.	.	.	1
122	.	.	.	2	.	.	.	.	.	.	1	.	.	.	.	.	.	.	.	.	.	.	.	.	4
117	.	.	.	3	1	.	.	1	.	3	.	1	2	3	.	.	.	.	.	.	.	.	.	.	12
112	.	1	1	4	3	2	1	.	.	3	3	4	2	1	1	2	1	.	.	.	.	.	.	.	26
107	.	1	1	4	5	.	.	3	1	4	1	4	2	1	1	1	1	1	.	.	.	.	.	.	31
102	2	2	2	3	4	1	1	1	3	3	9	9	8	3	1	1	1	1	.	.	.	.	.	.	53
97	.	3	1	4	4	8	3	1	5	11	9	6	15	8	3	5	1	6	1	.	.	.	.	.	94
92	3	1	.	4	6	5	2	7	3	15	11	14	11	5	8	1	3	.	1	.	.	.	.	.	100
87	1	3	.	4	7	5	3	8	5	8	13	16	12	9	4	3	4	3	1	.	.	.	.	.	109
82	1	1	.	1	1	2	.	1	3	16	8	8	6	2	5	9	.	3	.	.	.	.	.	.	67
77	.	3	.	4	.	4	2	4	3	8	4	5	5	5	6	3	.	1	1	.	.	.	.	.	58
72	1	.	1	3	.	.	.	1	1	6	3	5	3	1	1	2	.	2	1	.	.	.	.	.	38
67	.	.	1	1	.	.	1	1	1	3	3	7	3	1	1	.	1	1	1	.	.	.	.	1	24
62	.	.	1	1	.	.	.	1	1	3	4	4	1	1	2	.	1	.	.	.	.	.	.	.	16
57	.	.	.	2	1	.	.	.	.	3	3	2	1	2	.	2	.	2	.	.	.	.	.	.	13
52	.	.	.	.	.	.	.	.	.	3	3	1	1	1	.	.	.	.	.	.	.	.	.	.	10
47	.	.	.	.	.	.	1	.	.	3	1	1	1	.	1	.	.	.	.	.	.	.	.	.	5
42	.	.	.	.	.	.	.	1	.	2	.	.	2	.	1	.	.	.	.	.	.	.	.	.	5
37	.	.	.	.	.	.	.	.	.	1	.	.	.	.	1	.	.	.	.	.	.	.	.	.	2
32	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	.	.	.	.	.	.	.	.	.	1
27	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	2
22	.	1	.	.	.	.	.	.	.	1	.	.	.	.	.	1	.	.	.	.	.	.	.	.	3
17	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	2
12	.	.	.	.	1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0
7	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1
Total	8	16	7	37	34	30	14	31	25	91	75	87	75	44	43	29	11	18	6	0	0	0	0	1	682

simpleton. If these tendencies were regular the entries on Table IX should lie either along the horizontal line or the diagonal line indicated. As it is, however, the entries merely cluster around these two lines. The multitude of genetic and environmental factors which determine the M.R. of a given person blur the effect of segregation, especially when the two lines are close together.

The dual tendency of the scattering of mental ratios of relatives renders comparison of intelligence of relatives of different degrees difficult unless the grade of patient is specified. To facilitate such comparisons the patients were divided into two groups. In one group the M.R. was below 50 and in the other it was 50 or above; the reasons for subdivision at this point will be discussed later. The results are given in Table X. Patients with M.R. below 50 had relatives whose mean M.R. was 88.2 and those with M.R. above 50 had relatives with mean M.R. 85.8. The regression towards the normal is therefore much more marked in the case of the relatives of the first group of patients than in the case of the relatives of the second. If the patients' sibs (Table X) are first examined from this point of view, it is seen that, for those whose M.R. is 24.2, the sibs have a mean M.R. of 87.4; this mean has regressed at least four-fifths of the distance towards an M.R. 100 which might be presumed to be the norm for the general population in urban areas. The patients whose

Table X. *Stanford-Binet tests: mean mental ratios*

The patients are divided into groups according to mental ratio above or below 50; the relatives are classified according to degree of relationship.

Type of relationship	No. of pairs	Mental ratio of patient			Mental ratio of relative	
		Range	Mean	S.D.	Mean	S.D.
Female patient's child	14	0-49	44.7	3.4	76.2	14.1
	77	50 +	64.8	11.9	77.5	17.3
	91	Total	61.7	13.2	77.3	16.9
Patient's sib	120	0-49	24.2	12.8	87.4	19.5
	101	50 +	65.8	10.6	84.9	15.5
	221	Total	43.1	23.9	86.3	18.1
Patient's half-sib	16	0-49	30.3	11.4	93.8	17.3
	37	50 +	65.1	12.0	85.6	17.9
	53	Total	54.6	19.9	88.1	18.0
Patient's nephew or niece	74	0-49	33.9	11.5	95.4	16.9
	106	50 +	62.5	9.3	90.8	13.6
	180	Total	50.7	12.5	92.7	15.1
Patient's first cousin or equivalent relationship	69	0-49	42.3	11.6	82.8	17.2
	68	50 +	61.0	9.2	89.2	15.2
	137	Total	51.6	14.0	86.0	16.4
All types	293	0-49	32.2	14.0	88.2	18.7
	389	50 +	63.8	10.8	85.8	16.5
	682	Total	50.2	19.8	86.8	17.7

mean M.R. is 65.8, however, have sibs whose mean M.R. is 84.9. If mild mental impairment were inherited by multiple Mendelian additive factors, the sibs should regress towards the norm for the general population. On this hypothesis, the norm of M.R. for the general population from which these families have been drawn is 104.0.

If the group of patients' children is examined, the mean M.R. is seen to be almost the same whether the patients' own M.R. is below 50 or above. The tendency of the children to regress towards the normal is much less marked than in the case of sibs. If the mean M.R. of the children is, in fact, halfway between that of their mothers and that of the standard population, the implication would be that the population from which the fathers of these children were selected had a mean M.R. of 90.2. Inspection of the figures for mental ratios of mothers reveals that few of these patients have an M.R. below 50 and none below 38. The number of children born to people with I.Q. below 50 is very small, and a line drawn between 45 and 50 separates the infertile from the fertile group of defectives. It was in order to confine the analysis of regressions to the normal and the fertile group that the limit of M.R. 50 was decided upon (see Table XI).

Table XI. *Stanford-Binet tests: estimates for mean of general population based upon regressions of M.R. of relatives of patients with M.R. above 50*

Type of relationship	No. of pairs	Patients' mean mental ratio	Relatives' mean mental ratio	Estimate for mental ratio of general population
Female patient's child	77	64.8	77.5	90.2
Patient's sib	101	65.8	84.9	104.0
Patient's half-sib	37	65.1	85.6	95.8
Patient's nephew or niece	106	62.5	90.8	104.9
Patient's first cousin or equivalent relationship	68	61.0	89.2	96.2
All types	389	63.8	85.8	99.4

The half-sibs of patients are, on the average, less intelligent than the sibs but not greatly so. Half-sibs were most often to be found in families where there was illegitimacy and parents in such cases were not always up to the average for general intelligence. As usual, the relatives of idiots were more intelligent than the corresponding relatives of patients with M.R. above 50. In the case of half-sibs the regression should be three-fourths of the distance towards the normal, and the implication here is that patients with M.R. over 50 are drawn from a population whose mean M.R. is 95.8. The nephews and nieces of patients with M.R. above 50, on

the same hypothesis of multifactorial inheritance of mental ability, could have been drawn from a population with mean M.R. 104.9.

The group of patients' first cousins was probably not a fair sample of such relatives in general for it contained an inordinately large number of children from one rural district. The mean M.R. of the population from which the patients and their first cousins were selected can be estimated on the assumption that the mean intelligence of cousins of the higher grade patients regresses seven-eighths of the distance towards the standard norm. Here the standard M.R. would be 96.2.

A weighted average for all types of pairs of high-grade patient and relative shows that for 389 pairs the regression is in close agreement with expectation. For the group of patients with M.R. of 50 or over, the ability of relatives is, on the average, such as would be expected if the general population had a mean mental ratio of 99.4 and if the genetic background of mental ability were multifactorial.

IV. ANALYSIS OF PORTEUS MAZE TEST RESULTS

With one or two exceptions, the effect of testing patients and relatives with the Porteus Maze test gave rise to figures quite similar to those which were obtained by using the Binet tests. In Tables XII and XV they are set out in detail with relatives classified by degree of relationship and by sex. For the high-grade patients, Table XIII is complementary to

Table XII. *Porteus Maze tests: mean mental ratios*

The patients are divided into groups according to mental ratio above or below 50; the relatives are classified according to degree of relationship.

Type of relationship	No. of pairs	Mental ratio of patient			Mental ratio of relative	
		Range	Mean	S.D.	Mean	S.D.
Female patient's child	23	0-49	39.1	6.7	71.5	17.8
	30	50+	69.6	15.7	81.6	18.2
	53	Total	56.3	19.0	77.3	18.4
Patient's sib	81	0-49	24.6	13.5	86.2	25.7
	73	50+	73.8	16.7	86.6	19.9
	154	Total	47.9	27.2	86.4	24.1
Patient's half-sib	10	0-49	34.3	8.0	86.3	20.9
	23	50+	80.1	18.3	80.9	24.8
	33	Total	66.2	26.5	82.5	23.9
Patient's nephew or niece	41	0-49	30.8	10.6	94.4	21.4
	89	50+	73.6	19.2	93.7	18.4
	130	Total	60.1	26.1	93.9	19.3
Patient's first cousin or equivalent relationship	69	0-49	40.9	10.5	82.8	20.7
	42	50+	70.5	13.2	87.8	20.4
	111	Total	52.1	18.5	84.6	21.1
All types	224	0-49	32.6	13.4	85.2	23.3
	257	50+	73.3	17.3	88.1	21.0
	481	Total	54.5	25.5	86.7	22.2

Table XIII. *Porteus Maze tests: estimates for mean of general population based upon regressions of mental ratio of relatives of patients with mental ratio above 50*

Type of relationship	No. of pairs	Patients' mean mental ratio	Relatives' mean mental ratio	Estimate for mental ratio of general population
Female patient's child	30	69.6	81.6	93.6
Patient's sib	73	73.8	86.6	99.4
Patient's half-sib	23	80.4	80.9	81.3
Patient's nephew or niece	89	73.6	93.7	103.7
Patient's first cousin or equivalent relationship	42	70.5	87.8	92.1
All types	257	73.3	88.1	97.5

Table XIV. *Stanford-Binet tests: mean mental ratios*

The patients are divided into groups according to mental ratio above or below 50; the patients and their relatives are classified according to sex.

Type of relationship	No. of pairs	Mental ratio of patient			Mental ratio of relative	
		Range	Mean	S.D.	Mean	S.D.
Male patient's male relative	59	0-49	31.0	11.2	89.5	20.0
	64	50 +	63.9	10.8	89.1	17.5
	123	Total	48.2	19.6	89.3	18.2
Male patient's female relative	74	0-49	26.8	13.7	90.9	16.9
	72	50 +	62.2	8.9	90.1	17.3
	146	Total	44.2	21.3	90.5	17.1
Female patient's male relative	77	0-49	36.1	13.3	85.5	19.2
	142	50 +	64.1	11.5	83.0	16.8
	219	Total	54.3	18.0	83.9	17.6
Female patient's female relative	83	0-49	34.2	15.2	87.3	18.9
	111	50 +	64.3	10.5	84.7	13.8
	194	Total	51.5	19.4	85.8	16.2
All types	293	0-49	32.2	14.0	88.2	18.7
	389	50 +	63.8	10.8	85.8	16.5
	682	Total	50.2	19.8	86.8	17.7

Table XV. *Porteus Maze tests: mean mental ratios*

The patients are divided into groups according to mental ratio above or below 50; the patients and their relatives are classified according to sex.

Type of relationship	No. of pairs	Mental ratio of patient			Mental ratio of relative	
		Range	Mean	S.D.	Mean	S.D.
Male patient's male relative	39	0-49	29.7	10.1	95.9	20.5
	44	50 +	72.5	18.2	91.2	23.2
	83	Total	52.4	24.7	93.4	22.1
Male patient's female relative	56	0-49	24.2	11.6	87.2	24.8
	51	50 +	73.9	14.7	90.6	18.1
	107	Total	47.9	28.1	88.8	22.0
Female patient's male relative	55	0-49	38.9	13.3	87.1	20.6
	90	50 +	72.5	19.5	87.9	21.0
	145	Total	59.7	23.8	87.6	20.8
Female patient's female relative	74	0-49	36.0	13.2	76.4	22.6
	72	50 +	74.2	17.4	85.1	20.0
	146	Total	54.9	24.4	80.7	21.9
All types	224	0-49	32.6	13.4	85.2	23.3
	257	50 +	73.3	17.3	88.1	21.0
	481	Total	54.5	25.5	86.7	22.2

Table XI, and the estimate of the mean M.R. of the general population, based on the weighted mean of all the regressions, is 97.5. The general result points towards the conclusion that the genetic background of ability to perform the Porteus Maze test is similar to that of ability to perform the Binet test. In a small group of half-sibs, the relatives had scarcely any more ability than the patients. The patients in this group were drawn from rather a specially inferior social class selected mainly by illegitimacy. For sibs the regression towards the normal is almost perfect.

V. SEX DIFFERENCES

In order to test various hypotheses about the influence of sex on the inheritance of mental ability, each pair of scores of patient and relative was arranged in one of four classes according to sex, in Tables XIV and XV. When all grades of patient were taken together, no difference could be found which was attributable to the action of sex-linked genes. The male patients had relatives who scored higher than the relatives of the female patients, but this might be partly accounted for by the predominance of low-grade male patients. Moreover, in the Binet results, the female relatives were slightly, but consistently better, and in the Porteus results slightly worse, than the male relatives. In both cases, the effect was independent of the sex of the patient. If genes located on the sex chromosomes were responsible for major variation in intellectual capacity, the male relatives of male patients and female relatives of female patients should have been markedly inferior to the rest of the relatives tested. There is no evidence of such a phenomenon in the Binet or Porteus test results even if low or high grade patients are considered separately.

VI. CONCLUSION

The Stanford-Binet test was given to 682 relatives of mentally defective patients and the Porteus Maze test to 481 such relatives. A method of scoring mental ratio was devised which enabled comparison to be made between the abilities of patients who were mostly adults and their tested relatives who were mostly children. In summarizing the results, the patients were classified according to their sexes and divided into two groups according to whether M.R. was below 50 or not. Relatives were classified according to sex and mental ratio. Each pair of patient and relative was considered as a separate datum although some patients had more than one relative.

The results agree with those of the larger survey previously published, in which mental ability was only rated, in that it was found that the relatives of patients with M.R. below 50 had better scores on the whole than the relatives of patients with M.R. 50 or above. Analysis of the two groups of patients separately, showed that familial cases in the low-grade, infertile group were rare, but when they occurred were usually low-grade. The ability of relatives of patients of M.R. 50 or over tends to regress towards the normal as would be expected if the factors underlying inheritance of mental ability were multiple additive ones.

Certain irregularities in the results can be interpreted as due to the selection of some of the institutional patients from social strata in which mental ability is more often subnormal than usual. No evidence was found of effects which could have been due to genes on the sex chromosomes which determine mental ability.

The investigation was carried out under the auspices of the Medical Research Council and the Darwin Trust.

REFERENCES

- (1) PENROSE, L. S. (1938). "A clinical and genetic study of 1280 cases of mental defect." *Spec. Rep. Ser. Med. Res. Coun., Lond.*, no. 229. London: H.M. Stationery Office.
- (2) BURT, C. (1922). *Mental and Scholastic Tests*. London: P. S. King and Co.
- (3) HERRMANN, L. & HOGGEN, L. (1933). *Proc. Roy. Soc. Edinb.* LIII.
- (4) RICHARDSON, C. A. & STOKES, C. W. (1933). "Growth and variability of intelligence." *Brit. J. Psychol. Monogr. Suppl.*, no. 18.
- (5) MILES, C. C. & MILES, W. R. (1932). *Amer. J. Psychol.* XLIV, 44-78.

(*Manuscript received 7 October 1938*)

TWO CONTRIBUTIONS TO THE THEORY OF MENTAL TESTING

I. A NEW PERFORMANCE TEST FOR MEASURING ABILITIES AS CORRELATION COEFFICIENTS

BY W. STEPHENSON

From the Institute of Experimental Psychology, Oxford

- I. *Introduction* (pp. 19-20).
- II. *Premises* (pp. 20-23).
- III. *Principle of the new performance test population* (pp. 23-28).
- IV. *Description of the performance series* (pp. 28-30).
- V. *Performance times* (p. 31).
- VI. *Ability measured by correlation coefficients* (pp. 31-33).
- VII. *Conclusion* (pp. 33-35).
- References* (p. 35).

I. INTRODUCTION

THE present paper continues those in which I have offered an introduction to *Q*-technique^(1, 2), that is, the factor study of *persons* as variables instead of as items in a statistical population of individuals. In the present paper and one to follow it, I propose to describe two applications of *Q*-technique to mental testing, one of a particular nature concerning a novel way of measuring ability in terms of correlation coefficients, and the other of general and practical applicability, in which the various qualitative features of a person's responses in a performance test are used as a statistical 'population' for correlational purposes and factor analysis.

The setting for the first of these two applications, which is the concern of the present paper, is briefly as follows. Measurement of abilities has hitherto depended upon statistical populations of persons, either for seeking *average* performances (as in the case of standardization of a test like the Binet), or for supplying norms, or standardized scores. Now *Q*-technique has no necessary dependency upon any population of persons, and I have therefore asked myself whether ability might be measured in a systematic way in terms of only *one* person's ability, much as length is measured in terms of a standard yard. The thesis to be maintained in the sequel is to the effect that if a standard ability is taken,

such as that of say an Einstein, other abilities might be measured in terms of it as correlation coefficients. The performance test I have devised is a solution, however crude in nature, of the question I set myself.

This application, however, serves in addition an ulterior purpose, concerned with the general theory of mental testing and factor analysis. Work on the correlation of persons has been conceived by Burt⁽³⁾ as merely an alternative procedure to that of correlating tests or traits, it being a matter of convenience whether we correlate persons or tests. According to Burt, if there are more persons than tests, then the tests should be correlated, but if there are more tests than persons, then the persons should be correlated instead. In my opinion far more than mere convenience is involved in whether we correlate persons or tests. There are conditions under which it is necessary and sufficient to correlate persons, and other conditions when it is no less incumbent upon us to correlate tests. Some of these conditions are to be examined elsewhere⁽⁴⁾, but others of a particular nature deserve consideration, and I propose to draw attention to these in the course of the present paper and one to follow it.

II. PREMISES

Present-day mental tests are essentially *variables*, depending upon and varying with respect to a population of persons. There are obvious and essential differences between variables on the one hand, and statistical populations on the other; each involve assumptions peculiar to themselves. Thus, if I am working on verbal intelligence I do not dip my hand into a bag of verbal tests and select a number at random—I construct the tests I want, with an eye to certain psychological hypotheses that I wish to experiment upon. But I *do* select the necessary statistical sample of persons in some such random way. Under conditions of this kind nothing is gained, and no purpose or intention is changed, if I merely reverse the respective roles of persons and tests on grounds of expediency, even supposing for the moment that for the sake of convenience I could so change a variable into a unit of a statistical population, and units of a population into chameleon-like variables. The very verbal tests owe their *form* and method of construction to the exigencies of psychological theory on the one hand, and to the needs of a statistical regard of persons on the other. We must take care, therefore, to see that these purposes and principles are not just conveniently overlooked when we seek to correlate persons instead of tests, and tests instead of persons.

If the technique of correlating persons has to have its own place in

mental test science, or in psychometry in general, it must set its own conditions and be based on assumptions peculiar to itself. These must be of such a nature that it is necessary and sufficient to use persons as variables, and tests or qualities themselves as units of a statistical population. Moreover, the psychological hypotheses should relate specifically to the *persons*, whilst purely statistical conditions should govern the tests or qualities. It is for conditions and principles of this kind that I have looked in my regard of *Q*-technique.

In normal mental test theory there is a test variable (consisting, if need be, of several sub-tests which are used additively) in terms of which each person obtains a score; these scores are usually approximately normally distributed, and are standardized—that is reduced to values with mean 0, and standard deviation 1. One test is then correlated with another to see how far the two may be presumed to measure one and the same ability; or many tests are intercorrelated, with a very similar objective. These tests need satisfy no statistical conditions such as pertain to a universe from which samples are drawn. Indeed, on occasion there might be only *one* test of a particular kind in existence, and yet obviously it could satisfy all the necessary conditions of a variable with respect to a population of persons. Thus, two tests of completed and interrupted tasks, of the kind devised by Zeigarnik⁽⁵⁾, cannot be applied to one and the same group of persons—for as soon as they have performed the first test they know what to expect in the second, and this completely alters the nature of the latter. In effect only one test of this kind can ever exist for a given population of persons. Yet valuable psychological work might be done with this one test variable.

As for the statistical population of persons, its constituents are considered to be discrete and separate entities. But many assumptions enter into the grouping of human beings into a sample; it is in fact notoriously difficult to obtain a satisfactory statistical population anyhow or anywhere. Thus it is not permissible to include Negroes, Chinese and Whites into one person-population; nor would the psychologist care to mix children from London with those from Leeds. Physically the individuals are separate, but psychologically they are complexly related. Persons are bound together by ties of race, environment, education, training, social and civic culture, and if any of these are too firmly knotted for a section of human beings, that section cannot be included in the universe of a population from which samples are to be drawn. The essential criterion for a statistical universe is that it should be possible to regard all its constituents as dependent upon innumerable

small influences, interacting in infinite complexity, so that the conditions are produced for the theory of "errors" to apply. Such a theory can account for the normal distribution (or indeed for other distributions of this kind as well) of the quantities for which the statistical population of persons has been appraised or measured⁽⁶⁾. As is well known, both Spearman and Thomson have considered conditions of this kind in seeking to account for the substructure of abilities⁽⁷⁾.

Analogous conditions for ability measurement in terms of *Q*-technique would require us to contemplate *one* person, and *only* one if need be, as a variable, and a statistical population of tests of some kind from which samples could be drawn. For any such sample, each of the tests would obtain a score as a result of the individual's operation upon it; but these scores would not be added, any more than the scores for different persons are added when a population of persons is measured; instead they would have a more or less normal distribution in terms of some common qualitative and quantitative character. Their mean for this character would be zero (when standardized), and their standard deviation 1.

Even up to this point in the argument there exist no tests anywhere that would satisfy these conditions, unless only *one* kind of test is considered, involving one quality of performance, so that a common basis of quantification could be employed.

Moreover, for our special problem certain other conditions have to be satisfied. If one person is to be used as a standard in terms of whom to measure the ability of other persons by correlating these with the standard person's array; and if the correlations themselves are to be direct measures of these persons' ability, then the population of tests has to be of such a nature that only persons as brilliant as the standard can gain similar scores, test for test. Others less capable have to gain different scores for these same tests, test for test; and stupid persons have to obtain either random scores (relative to those of the standard person), or else turn tables upon the scores obtained by the capable ones—in tests for which the clever persons obtain a high score the stupid ones have to gain a low score, and vice versa. When the arrays of scores for the various persons are correlated with the standard array, only those persons as brilliant as the standard will correlate highly with him, and those less capable will correlate less with him, and less in proportion to their lesser abilities. Certainly no published mental tests known to me could be thrown together to constitute a population of tests fulfilling these conditions! It is obvious, too, that if such a test could be made it would have to be of a form unlike any at present known.

If it were indeed possible to construct such a population of tests it might well be asked *why* we should go to all the trouble of measuring ability in this way, as a correlation coefficient, instead of by the usual way. It is possible, of course, for us to set up all kinds of bizarre standards in terms of which to measure individuals. I certainly did not think of making any such purely arbitrary scales. Instead I hoped to conceive of a test with the same kind of systematic method and procedure attached to it that we expect in a science of mental testing. The new form of test, for example, has itself to supply individual differences in an orderly way, but *without any* necessary dependency upon statistical populations or regard of persons. To achieve this is at the moment a matter of academic interest only, but it is important to pay regard to the possibilities. So much is taken for granted in current mental test theory that everything is to be gained by questioning the assumptions involved, and perhaps still more by experimenting with others. And this, indeed, is the essential purpose of the new test. It will serve to demonstrate that *Q*-technique can have its own special problems, and that its assumptions, principles, and objectives can be peculiar to itself, and independent of those involved in present-day testing theory. For the present it is immaterial whether the test is practical or impractical; it is merely incumbent upon us to think along any new lines that *Q*-technique may suggest to us.

III. PRINCIPLE OF THE NEW PERFORMANCE TEST POPULATION

Many test populations of the required kind could be devised. But for the purpose of the second paper I need some performance tests, and I have therefore made the following test of a performance kind (I shall call it as a whole a *test*, but it is really conceived as a sample of tasks, drawn from a universe of similar tasks that could be made). It will be remembered that an illustrative rather than a practical matter of mental testing is under consideration, and it is in no way intended that the following test should be used for work-a-day measurement of ability.

The performance test has the following construction. It consists of twenty tasks, each like a large-scale jig-saw puzzle. The individual has to make twenty copies of a standard shape from given pieces; these copies are not exact duplicates of the standard, but may diverge considerably from it (see Fig. 1). Having constructed the twenty copies under time-limit conditions *the individual is asked to appraise them for their degree of likeness to the standard. All the copies have this character in common, i.e. some degree of similarity to the standard; it is this that affords a basis for their quantification.*

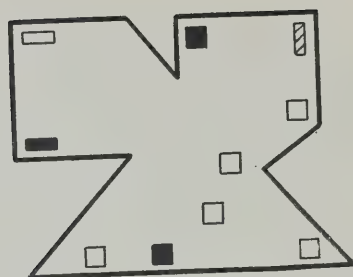
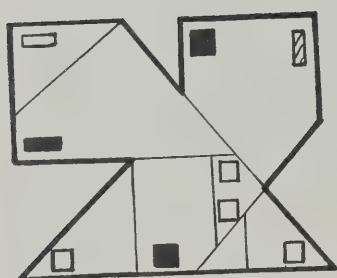
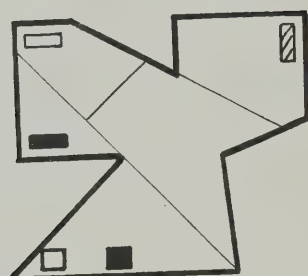


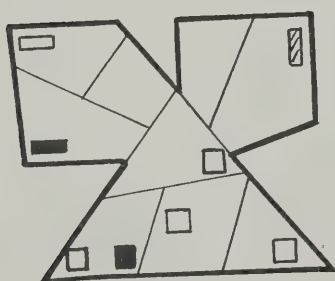
Fig. 1. STANDARD.



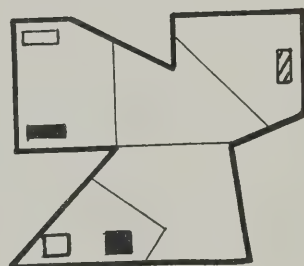
A



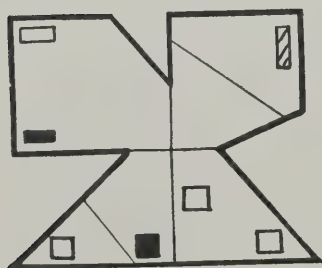
D



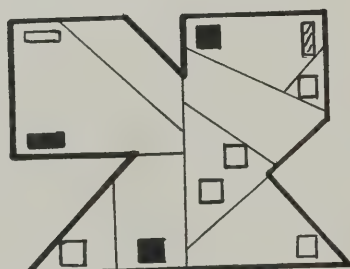
B



E

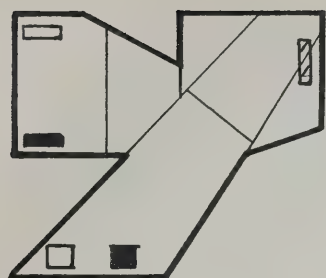


C

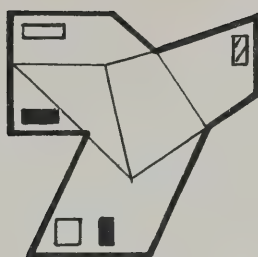


F

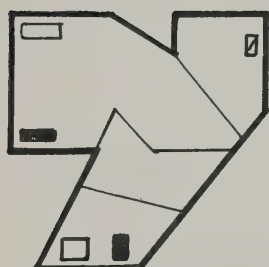
Fig. 1. A—F.



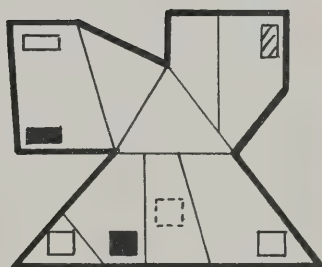
G



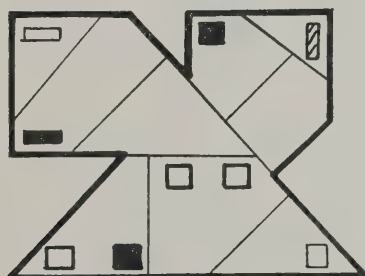
K



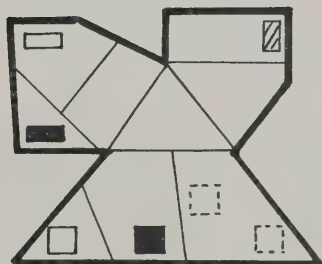
H



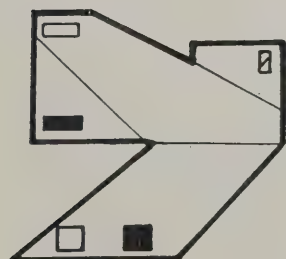
L



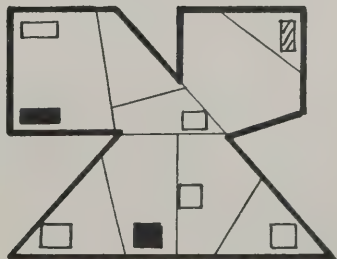
I



M



J



N

Fig. 1. G-N.

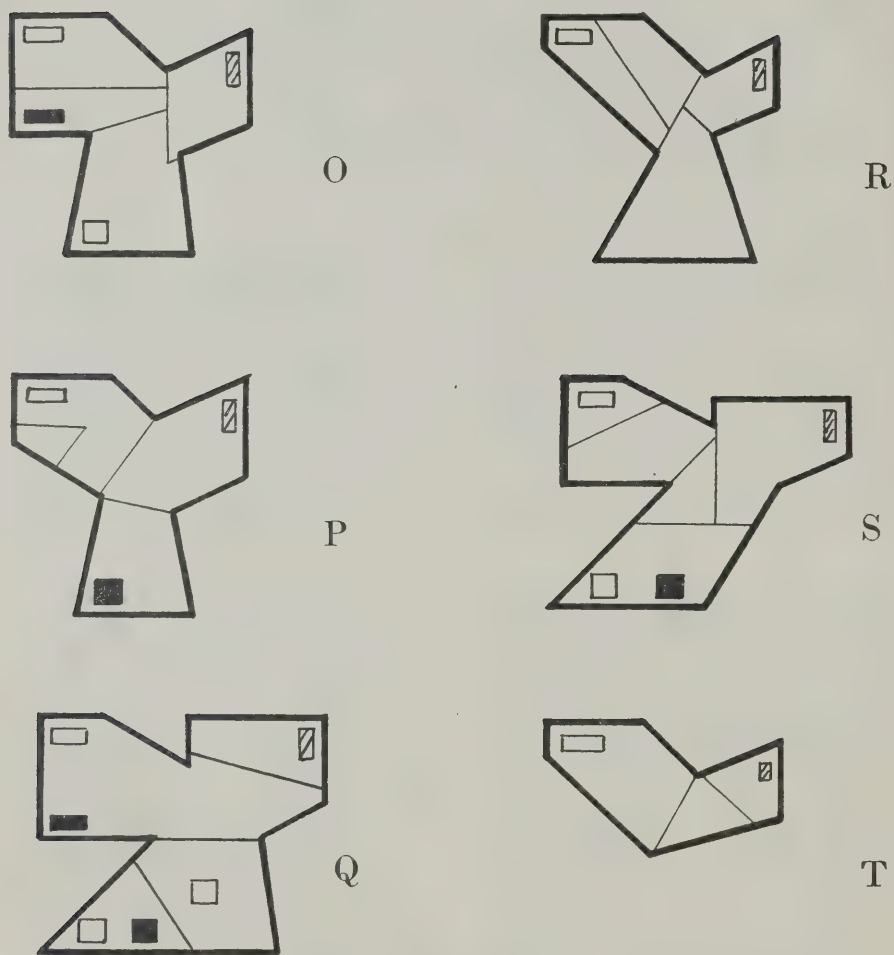


Fig. 1. O-T.

Fig. 1. Showing the standard shape and the 20 completed copies A to T, in order of production. Faint lines mark off pieces composing the copy; thus A has 8 pieces, and B to T, 9, 6, 5, 5, 9, 6, 4, 9, 4, 6, 9, 9, 9, 4, 4, 4, 4, 5 and 3 pieces respectively. The copies depart from the standard in the order F, A, I, B, N, C, L, M, Q, D, E, G, S, J, H, K, O, P, R, T. Full black squares or rectangles are *red*, empty ones *blue*, hachured ones *green*, and dotted ones *bright (light) green*. The latter colour is not on the standard. All drawings are to scale 1:7.

Now only twenty tasks are involved. But two or three hundred could be made if need be, and for a large number like this the appraisal for similarity to the standard would supply a frequency distribution of the quantities, and the distribution could theoretically and practically be approximately normal in shape. Only a few of the productions would be very similar to the standard; only a few very little like it; and most more or less like it. (There are other conditions, as we shall see, that would ensure this.) For my small sample of twenty tasks only, however, I shall use a rank order. The copy most like the standard will be placed first in the rank, and that least like it last, with the others duly in between in order.

Now if the twenty tasks are correctly made there is a correct order in which they can be placed when appraised for similarity to the standard. But if the individual makes errors in constructing his copies, or does not finish them in time, the result will be that he cannot supply the correct rank order. Moreover, only really capable persons can solve the tasks in the given time, and therefore they alone have the opportunity of ranking their productions correctly; and the less capable the person the more upset is his rank order relative to the standard, correct one.

This comes about in several ways. Persons of poor ability cannot construct the more difficult copies of the standard in the allotted time, and the more difficult the task the more in fact it resembles the standard model. Thus when such a person ranks his productions those that ought to have been placed high up in the rank order are of necessity placed lowly; and the less capable the person the more lowly will he place the tasks which are, when correctly solved, most like the standard model. On the other hand, these persons usually manage to construct certain easy copies, which are superficially very similar to the standard, and these are therefore very readily placed first in their rank order! Similarly throughout the sample of tasks there are sources of error which result in displacements of this kind, with the net result that the more stupid the person the more he tends to place the *easiest* tasks (which are really least of all like the standard) at the top of his rank order.

Now when the rank orders of different persons are correlated the desired result is obtained: clever individuals correlate highly with the standard rank order, and dull ones correlate less highly with it, and the more dull they are the less highly do they correlate with the standard array. Compared with existing forms of mental tests this one seems to turn things completely inside out, like a sausage with its skin on the inside. But the test as a whole does what is required, as the following fuller description of it, and a set of results obtained by its use, will show.

I would again remind the reader that although only twenty tasks are involved in the present sample a very much larger sample could be used if need be. Any such large sample would have its own *correct* frequency distribution in terms of which to compare those supplied by each individual for the appraisal of his productions for their likeness to the standard model.

IV. DESCRIPTION OF THE PERFORMANCE SERIES

The standard model is placed in front of the subject to be tested. Twenty copies of this have to be made by him from twenty sets of pieces. Any one set consists of from three to nine pieces. Only one of the sets provides, if it is correctly put together, an exact copy of the standard model, the rest diverging from it progressively, and the subject is asked to make as exact a copy of the standard as he can with the given pieces of each set. The copies that he *can* make are objectively determined by the given pieces, and if the subject works deliberately and is prepared to construct his copies somewhat imaginatively it is possible for him to arrive at a conclusion satisfactory to himself and to the test constructor. The twenty copies he is expected to make are shown in Fig. 1 (A to T). The pieces (one set at a time) are given to the subject in random order, with reverse sides uppermost; they have to be turned over and placed in juxtaposition, much as dominoes are placed side by side on a table, in such a way as to provide the required copy of the standard.

All pieces are made of $\frac{1}{4}$ in. soft pine, varnished and numbered on the reverse side, painted white along all edges, and enamelled a bright lavender colour on the face. Sizes are indicated in Fig. 1, which shows the pieces drawn in proportion, in the positions they take up when the copies are 'correctly' put together. The standard and various pieces of most sets have painted upon them a number of coloured patches, red and blue squares, and red, blue and green rectangles of a regular size throughout. These, and the shapes of the pieces, offer the clues that are needed for solving the copies.

The order of application of the sets is as shown in Fig. 1 from A to T: it will be observed that the copies diverge progressively from the standard (although not in the order of application of the sets) as we pass from A to T: the first few are not unlike it, but those towards the end of the series are only abstractly like it. The copies actually diverge in a logical way according to a very simple principle of decretion—colours are omitted and sections cut away from the standard shape; this logical order supplies the series given in the footnote to Fig. 1. After performing the first few sets the subject is expected to have learnt that the copies he can make fall into some such series.

The subject sits at a small table with the standard model in front of him, inclined on a book-rest at the back of the table. A smooth-surfaced strawboard (12×10 in.) is placed on the table in front, having upon it the pieces of a set, all of which, to begin with, are face downward, each piece having a separate (random) position. The following instructions are given before the individual begins:

"These pieces (set A), when turned over (demonstrate with one piece) can make a copy of this standard model (point). This set of pieces makes almost an exact copy of the standard when the pieces are correctly placed side by side; but the later sets do not make exact copies of the standard. You will be asked to make twenty copies, from twenty different sets of pieces, of which this set is the first.

"The first few sets make copies that do not differ much from the standard, but the later sets may differ considerably from the standard. Nevertheless, if you follow all the clues given on the pieces you can nearly always construct a model something like the standard. In each case you have to construct a copy which will be as nearly like the standard as it is possible to make it with the given pieces.

"With these pieces, then, see how quickly you can make a copy of this standard, as nearly as possible like the standard, using all the given pieces."

The individual's construction is left on its strawboard base, and is lifted bodily, without disturbing it, on to a low table at his side. The next set of pieces is then placed before him, and the testing proceeds until all twenty copies have been attempted. The time taken to construct a copy is recorded by the tester; and if a copy is not finished at the expiry of five minutes the subject is stopped from working at it. It is placed along with the completed copies in the state of construction reached by the individual.

Once they are correctly completed, the copies only rarely fail to satisfy the individual that they are the required ones. This is at first sight not an obvious feature of the test, because in most performance tests the pieces fit into recesses and this provides an automatic check on the correctness or otherwise of a production, whereas in the present case there are no recesses and the copies themselves are somewhat imaginatively conceived. The individual has to decide for himself, sometimes very explicitly, whether or not his productions 'will do'. He is helped in this by the 'wholeness' of the finished copy, and by the objective fact that the pieces 'fit together nicely'. The need to make a decision about a copy, however, is very often apparent. Otherwise the test, so far, is like any other of a performance kind, except perhaps that it embraces a more varied sample of intelligent activities—it entices the individual into many avenues of performance and thought. The test seems to involve, amongst other things, a certain need for inventive or imaginative use of material, orderly planning, making decisions, analysis and synthesis, space-sense (8), generalization, learning by experience, and the ability to execute a lengthy series of interconnected constructions. These qualities make it, on its own account, a test of no mean difficulty.

As in the case of any other test of this kind, the *time* it takes the individual to perform each task, or the number he successfully constructs in the allotted time, may be used as a measure of ability.

The second, and for our present purpose the most interesting, part of the testing now takes place. When all the sets have been attempted, the subject is asked to appraise them for their similarity to the standard (which he can carry about in his hand, and look at as often as he likes). That most like the standard has to be placed first, that next, second, and so on in rank order. In the case of unfinished constructions the individual is told that they are probably as satisfactory as any that could be made with the given pieces, and that he should judge them in that light.

As I have already indicated, only one copy can be made which is a complete duplicate of the standard. This is copy F in Fig. 1. The other copies diverge from the standard either with respect to shape, colours, or size. Thus, copies A and B are identical for shape and size, but one lacks a coloured square that the other has got, and it is to that extent logically less like the standard. There are differences of this kind between all successive copies, in terms of which one can decide objectively, at least on logical grounds, that one construction is more like the standard than another. This is particularly the case, of course, when incomplete constructions are involved—i.e. those that the subject has been unable to complete in the given time. Needless to say, judgements are rarely based on logical grounds alone—on the contrary most individuals seem to appraise the copies by ‘general impressions’, or even on aesthetic grounds. Only the most careful, and capable, subjects seem to appraise the copies rationally; and the more stupid the individual the more likely he is to misplace the copies due to inadequate analysis. Copies C, D, and E are very apt to be placed higher up the rank order on this account than their logical analysis would warrant.

I have available the results for forty students to whom I applied the new performance test, and these data should be sufficient to illustrate the test’s uses. The students were of an age, within two years, and were all drawn from various Colleges at Oxford. For each we have (i) the time taken to complete each of the twenty copies, with a record of errors and unplaced pieces, and (ii) his rank order for these twenty copies, appraised for their similarity to the standard.

V. PERFORMANCE TIMES

I propose to use the time taken by the subject to complete his copies (with a correction for errors or unplaced pieces) as a straightforward measure of ability in the test. That is, the person who can perform the tasks quickest will be considered to be the highest in the scale of ability in this test. This measurement can then be used as a basis of comparison with ability as measured in terms of a correlation coefficient. The time measure is therefore only of indirect interest to us, but a few details are necessary.

Table I shows the arithmetic means of the times taken by the forty students to solve or attempt each of the twenty copies A to T. Sets A and B take the longest time, and D and H the least. There is a tendency for the time, and the difficulty of the tasks, to decrease as we pass from A to T. Account has to be taken of errors and incomplete solutions, and a preliminary study indicated that it would be fair to add 50 sec. to the subject's time at a copy, for each piece left unplaced or incorrectly placed. With this correction the mean times per task are those given in column (a) of Table I. These corrected times will be used as straightforward measures of ability. In point of fact the test calls for high ability, and a study with other subjects has shown that these corrected times correlate -0.738 ± 0.076 (S.E.) with the results of a pencil and paper intelligence test. The reliability of the test, gauged by the Spearman-Brown formula in terms of these time scores, is 0.893 for the forty students.

Table I. *Showing average time for forty students to perform each item. Column (a) gives the same averages with a correction added for unplaced and misplaced pieces. (All times to the nearest second)*

Test item	Average time	(a) Corrected average	Test item	Average time	(a) Corrected average
A	230	296	K	100	113
B	231	278	L	215	292
C	148	177	M	192	229
D	52	58	N	201	267
E	110	120	O	95	105
F	157	183	P	66	74
G	132	140	Q	119	127
H	49	54	R	116	148
I	135	146	S	89	93
J	106	114	T	73	73

VI. ABILITY MEASURED BY CORRELATION COEFFICIENTS

Our crucial concern is with each subject's appraisal of his productions for their similarity to the standard model. This similarity supplies the needful quantitative basis in terms of which individuals can be cor-

related in turn with the rank order given by an expert performer to his own productions, and it is anticipated from the nature of the test as a whole that these correlations will be in proportion to each person's ability, approximately, as this is assessed by more usual methods.

That is, we require an expert to appraise his own productions; and the rank order that he supplies will thereafter be used as a standard array in terms of which any other person's ability can be measured. He will set an upper limit to the scale of ability, but this will be no practical disadvantage if the expert is sufficiently competent. All other persons will be scaled downwards from this standard array, for no one can correlate more than 1.00 with it, although many might correlate perfectly with it. In the present case it seems that the logical order of decretion, upon which the test was originally constructed, would not differ greatly from the rank order that an expert analytical thinker and performer would supply, and I therefore propose to use this order as the standard array. The order in question is that given in the footnote to Fig. 1.

Each student's rank array was therefore correlated (rank coefficients) with this standard array, and the results are shown in Table II. The correlations range from 0.99 to 0.52, with a mean 0.870, there being some skew towards perfect correlation. Lower correlations would result all

Table II. *Showing frequency distribution of rank correlations. The correlations are between the logically correct order and the subject's own order for his productions*

Correlation	No. of students supply- ing this correlation
1.000 to 0.990	1
0.989 to 0.980	2
0.979 to 0.970	1
.....	
0.969 to 0.950	4
0.949 to 0.930	6
0.929 to 0.910	3
0.909 to 0.890	5
0.889 to 0.870	4
0.869 to 0.850	3
.....	
0.849 to 0.810	3
0.809 to 0.770	3
0.769 to 0.730	1
0.729 to 0.690	1
0.689 to 0.650	—
0.649 to 0.610	2
0.609 to 0.570	—
0.569 to 0.530	—
0.529 to 0.490	1
Total 40 students	

round if the time allowed per copy were less than the five minutes allowed as maximum in the case of these students. Under the conditions obtaining for these forty students, all of whom were men, one or two women have supplied a correlation as low as 0.33 with the standard array.

It is these correlations that we hope will be proportional, approximately, to ability as measured in the more usual way. An indication of their success in this direction is afforded by using them as scores, and seeing how far they correlate with, say, the time-scores referred to in the previous section. The two arrays for the forty subjects (their time-scores for producing the copies in the first place, and their correlation with the standard array just mentioned) correlate -0.768 ± 0.044 (S.E.) (product-moment correlation). That is, the person who performs the twenty tasks in the *quickest* time (corrected for errors) tends to have the *highest* correlation with the standard array. To this extent, therefore, our original anticipations are fulfilled; and indeed it is quite obvious as one sees the performances and subsequent appraisals, that only the more competent subjects can supply a rank order that approximates adequately to the logically correct order. The correlation -0.768 is linear, so that within the limits of this sample of twenty tasks the correlational measures of ability are also linear. Otherwise they would seem to involve the usual conditions of any correlational surfaces.

VII. CONCLUSION

It has been shown, then, that a sample of tests can be made in terms of which one individual at a time can be correlated with the *ability* of another person, whose expertness is accepted as a basis of measurement, and that the correlations so obtained are themselves a direct measure of ability—at least to the extent shown by a correlation of 0.768 with ability as measured in the usual way by the *time* it takes an individual to solve the same tasks. The time-scores, and the appraisal arrays, reach very similar ends.

Several matters of theoretical interest are involved, and it was to illustrate these that the test has been made. It has been my object to demonstrate that whether we correlate persons, or tests, should not be thought of as merely two interchangeable ways of treating one and the same matrix of data. If *Q*-technique, the use of persons as variables, is to be of any interest and importance on its own account, it must seek new fields of endeavour in psychometry. Beginning from the fact that

modern mental test theory and practice is based on the use of tests as *variables* and persons as a statistical population, I have asked myself whether these roles might be reversed. Prof. Burt (3) has suggested one way of effecting this, but it is scarcely the kind of solution or point of view that I have in mind. It leads to no new principles or underlying premisses in mental test theory. According to Burt it is merely a matter of convenience whether we correlate persons or tests. But the purpose and function of a door is not changed by hanging it back to front; and what I want of the concept of correlating persons is a new door, or better a window, and not a badly hung back-to-front door.

The use to which I have put *Q*-technique in the present paper accepts the conditions of a person as variable: if need be, for instance, we could vary *one* individual, testing him when he is drunk, sober, drugged, or depressed, and determining the influence of these conditions on his ability as measured relative to a standard ability. The study also accepts the implications of a statistical population and regard of *tests*. It is not sufficient, in this connection, to dip into a catalogue of tests for our sample populations. Statistical populations of use to psychology are not so much random, as infinitely complexly related in their components. The sample of tasks used above has been based on such considerations; a number of productions have to be appraised for their similarity to a standard model, and it is theoretically possible to construct a universe of such productions, distributed normally for their degree of likeness to the standard model. The small sample used here is but a faint echo of this universe, of course. Under these conditions I have next asked whether it is possible to dispense with individual differences as the *basis* of measurement of abilities, and whether the use of statistical populations of persons is essential for this measurement. No statistical population of persons, or gross averaging of samples drawn from such populations, is anywhere involved in the theory of the above test population. It is *sufficient* to correlate persons one by one with the known standard in order to arrive at a methodical measurement of their abilities. The individual differences between persons issue from the testing as a consequence of the logical structure of the test material, and not because of any dependency on, or statistical formulations about, persons. In short, a new principle of systematic quantification of abilities is contemplated. Finally, to arrive at the necessary individual differences a new *form* of test construction has had to be devised, which makes use of principles (or tricks if you will) which are specific to the needs of a *Q*-technique approach to the measurement of ability. It is not suggested for a moment

that we should in practice seek to measure abilities as correlation coefficients: but if we were driven to it, it could be done.

In all this, *Q*-technique is reaching out towards its own place in the sun, towards its own particular functions in psychometry. It searches for conditions under which the correlational regard of persons, as variables, can have a unique place in mental test theory and practice. There is nothing merely permissive, no matter of convenience only, in this regard of persons who are correlated—on the contrary the very foundations of mental test theory are uprooted and examined, and fresh young trees are planted in their place.

I shall discuss other matters of theoretical interest pertaining to the above performance test series in a subsequent paper.

REFERENCES

- (1) STEPHENSON, W. (1936). "The inverted factor technique." *Brit. J. Psychol.* XXVI, 344.
- (2) — (1936). "The foundations of psychometry: four factor systems." *Psychometrika*, I, 195.
- (3) BURT, C. (1938). "Correlating persons." *Brit. J. Psychol.* XXVIII, 72 (see also *Brit. J. Med. Psychol.* XVII, 195).
- (4) STEPHENSON, W. *Type Analysis*. (Not yet published.)
- (5) ZEIGARNIK, B. (1927). "Das Behalten erledigter und unerledigter Handlungen." *Psychol. Forsch.* IX, 1.
- (6) WHITTAKER, E. T. and ROBINSON, G. (1932). *The Calculus of Observations*, Ch. VIII. London: Blackie.
- (7) SPEARMAN, C. (1928). "The substructure of the mind." *Brit. J. Psychol.* XVIII, 249.
- (8) KOUSSY, A. A. el. (1935). "The visual perception of space." *Brit. J. Psychol. Monog. Sup.* VII, No. 20.

(Manuscript received 18 September 1938)

INDIVIDUAL DIFFERENCES IN AMPLITUDES OF OSCILLATION AND THEIR CONNECTION WITH STEADINESS OF CHARACTER

BY ROBERT D. WALTON

From the Department of Psychology, University of Glasgow

- I. *Introduction* (p. 36).
- II. *Procedure* (pp. 36-37).
- III. *Frequency distributions* (pp. 37-38).
- IV. *Reliabilities* (p. 38).
- V. *Age and intelligence* (pp. 39-40).
- VI. *Intercorrelations and a group factor* (p. 40).
- VII. *Factorial analysis of intercorrelations* (pp. 40-41).
- VIII. *Estimation of steadiness* (pp. 41-42).
- IX. *Estimates and oscillation scores* (pp. 42-43).
- X. *Effect of mistakes* (pp. 43-44).
- XI. *Test loadings with the group factor for oscillation* (pp. 44-45).
- XII. *Summary* (p. 46).
- References* (p. 46).

I. INTRODUCTION

In a previous investigation⁽⁵⁾ it had been concluded tentatively that the amplitude of a person's oscillations in efficiency during a short test was associated with his steadiness of character. The results were however diminished in value by the fact that the methods of obtaining some of the data did not permit of a statistical treatment which would give trustworthy estimates of significance.

The present investigation has improved on the former by having

- (1) a larger number of subjects,
- (2) a wider range of tests,
- (3) statistical methods giving estimates of the significance of the results.

II. PROCEDURE

Ninety subjects were available instead of fifty-five, and seven tests were used instead of four. These included the previous four tests, viz.:

- (1) Crossing out digit pairs with sum 10.
- (2) Crossing out digit pairs with difference 3.

- (3) Writing letters of the alphabet.
- (4) Writing cardinal numbers backwards.

And also

- (5) 'Code' test, replacing each letter on a sheet by the one two places before it in the alphabet.
- (6) Crossing out pairs of consecutive digits.
- (7) Crossing out permutations of the same 4 digits.

Each test was applied twice, but only one testing per day was done with any one class. The procedure in giving the tests was the same as before. In tests (1), (2), (6), (7) the subjects had to cross out by a stroke \ the specified groups, and by a stroke / the other groups. Correction by a second stroke was allowed, giving $\times$, but this occurred rarely with most subjects. The specified pairs were scattered by chance drawing among the others in the ratio of 1 to 4, to avoid any rhythmic effects. The other three tests were of the 'production' type, involving writing on the part of subject. In each test 'Stroke' was called every 15 sec., and a vertical stroke was made to mark the place reached on the paper. Twenty periods were taken for each test, giving a total time of 5 min.

The measure of the amplitude of oscillation adopted was the same as previously, viz.:

$$\text{Amplitude "0"} = \frac{\text{Standard deviation of scores in 20 periods}}{\text{Mean score in 20 periods}} = \frac{\sigma}{M}.$$

Fatigue effects tending to lower the score towards the end of test were again found to be absent (see (5), p. 184) and no adjustments were made on that account.

The reliabilities of the tests were obtained by calculating the correlation coefficients between the two applications.

III. FREQUENCY DISTRIBUTIONS

The frequency distributions of the coefficients for amplitude of oscillation (σ/M) were investigated diagrammatically, and also by comparison of the observed frequencies with those to be expected from a normal distribution. The diagrams showed the familiar 'cocked hat' shape which is usually considered sufficient warrant for the use of the product-moment method of calculating correlations and their probable errors. The coefficient σ/M is not subject to any particular limiting value. The observed and calculated frequencies are given in Table I.

The observed frequencies are not widely different from the normal frequencies. Using the χ^2 criterion⁽²⁾ to examine the amount of the

38 *Individual differences in amplitudes of oscillation*

difference, it is found that none of the distributions is significantly different from a normal distribution.¹

Table I

	Range of σ/M					
	$M' - 3\sigma'$	$M' - 2\sigma'$	$M' - \sigma'$	$M' + \sigma'$	$M' + 2\sigma'$	$M' + 3\sigma'$
Calculated frequencies	1	11.5	27	27	11.5	1
Observed frequencies: Test 1	0	15	31	20	10	3
Test 2	1	10	34	17	14	3
Test 3	2	13	29	23	7	5
Test 4	1	17	25	24	9	3
Test 5	1	9	35	19	13	2
Test 6	2	16	17	29	13	2
Test 7	0	11	32	28	4	4

M' and σ' refer to the distribution of σ/M .

The frequency distributions for Ages, i.q.'s, and Mean Scores were also investigated diagrammatically, and the corresponding figures, with expected values, are given in Table II. These three distributions are more nearly normal than the previous ones. The age distribution is approximately normal because the classes available happened to be one first form, two second forms, and one third form.

Table II

	Range					
	$M - 3\sigma$	$M - 2\sigma$	$M - \sigma$	$M + \sigma$	$M + 2\sigma$	$M + 3\sigma$
Calculated frequencies	2	11.4	28.6	28.6	11.4	2
Observed frequencies (age)	2	15	26	27	12	2
Observed frequencies (mean score)	4	12	26	28	9	5
Calculated frequency (i.q.)	1.7	10.6	26.2	26.2	10.6	1.7
Observed frequency (i.q.)	3	14	24	21	10	5

IV. RELIABILITIES

The above considerations justify the use of the product-moment method of calculating reliabilities and inter-test correlations. The reliability correlations are given in Table III. The coefficients vary between 4 and 10 times their probable error, and must therefore be taken as significant.

Table III

	Test						
	1	2	3	4	5	6	7
Correlation between 1st and 2nd applications	0.48	0.54	0.36	0.475	0.39	0.56	0.39
P.E. of r	0.06	0.06	0.08	0.06	0.07	0.05	0.07

¹ In applying the χ^2 criterion, columns 1 and 2 were taken together, and columns 5 and 6, giving four classes only.

V. AGE AND INTELLIGENCE

In the previous investigation no account was taken of intelligence, as I.Q.'s were not available. For the present tests I.Q.'s were known for most of the subjects.¹ The mean values of σ/M from the two applications were found for each test, and the correlations between each test and I.Q. calculated. The correlations are given in Table IV.

Table IV

	Test						
	1	2	3	4	5	6	7
r (with I.Q.)	0.06	0.04	-0.12	0.09	-0.04	-0.005	0.08
P.E. of r	0.08	0.08	0.08	0.08	0.08	0.08	0.08

Mean value of $r = +0.015$.

Hence it seems reasonable to assume that the measure of oscillation amplitude used is almost independent of intelligence level. In corroboration of this, the correlation between the Mean Score for all the tests, and I.Q. was $r = 0.012$ (P.E. 0.078).

The connexion between Mental Age, as found from the intelligence test results, and the oscillation scores was also investigated, giving the correlations shown in Table V. The correlations for tests (1), (3), (5) are significant. The correlation between Mental Age and the Mean Score from all the tests was $r = -0.284$, P.E. = 0.074.

Table V

	Test						
	1	2	3	4	5	6	7
r (with Mental Age)	-0.27	-0.08	-0.29	-0.06	-0.42	+0.06	-0.04
P.E. of r	0.07	0.08	0.07	0.08	0.07	0.08	0.08

The correlations between age and the mean score for each test were found to be as shown in Table VI. The correlations for tests (1) and (3) are significant, and for (4), (5) and (7) probably significant (i.e. probability of chance occurrence is $< 0.1(2)$). They tend to confirm the result suggested in the previous investigation of an inverse relationship between oscillation and age for the range of ages tested. Previously the correlations for writing digits (test 4) and writing letters (test 3) were insignificant. One is now significant ($r = 4$ times P.E.) and the other probably significant ($r = 2\frac{1}{2}$ times P.E.).

The correlation between the Mean Score in all the tests and age is $r = -0.295$, P.E. 0.067, which is significant.

¹ The Simplex Group Test was used. The average I.Q. was about 109.7. The average for the school as a whole was, on several occasions, found to be almost exactly 100, and the distribution very close to normal. The present subjects were from classes taking a full academic course, and hence above the general average.

Table VI

	Test						
	1	2	3	4	5	6	7
r (with age)	-0.37	-0.065	-0.29	-0.19	-0.21	+0.026	-0.18
P.E. of r	0.06	0.07	0.07	0.07	0.07	0.07	0.07

VI. INTERCORRELATIONS AND A GROUP FACTOR

The intercorrelations of the seven tests, using the mean value of σ/M for each, were found for those subjects, numbering seventy-nine, who had completed the fourteen applications. The twenty-one coefficients obtained were all positive (previously some were negative but not significant). Fourteen of them were significant ($r=3\frac{1}{2}$ times P.E.) and the lowest of the other seven was $1\frac{1}{2}$ times its P.E. When these coefficients were corrected for attenuation, using the reliability figures from Table III and the formula

$$r_{AB} = \frac{\sqrt{(r_{A_1B_2} \cdot r_{A_2B_1})}}{\sqrt{(r_{A_1A_2} \cdot r_{B_1B_2})}},$$

the values ranged from 0.33 to 0.90.

The intercorrelations were also found after the effect of age had been eliminated by the use of the usual formula for partial correlations. The figures obtained were little different from the raw coefficients, and were still all positive, with thirteen significant. The values are given in Table VII where the raw correlations are in the lower half, the age corrected figures in the top half in italics.

Table VII

	Test						
	Adding digits	Subtracting digits	Writing letters	Writing numbers	Code	Consecutive digits	Rearranging digits
Adding digits		<i>+0.288</i>	<i>+0.314</i>	<i>+0.206</i>	<i>+0.185</i>	<i>+0.401</i>	<i>+0.306</i>
Subtracting digits	<i>+0.291</i>		<i>+0.205</i>	<i>+0.213</i>	<i>+0.366</i>	<i>+0.341</i>	<i>+0.108</i>
Writing letters	<i>+0.388</i>	<i>+0.215</i>		<i>+0.426</i>	<i>+0.510</i>	<i>+0.206</i>	<i>+0.172</i>
Writing numbers	<i>+0.258</i>	<i>+0.221</i>	<i>+0.456</i>		<i>+0.281</i>	<i>+0.343</i>	<i>+0.207</i>
Code	<i>+0.244</i>	<i>+0.371</i>	<i>+0.538</i>	<i>+0.302</i>		<i>+0.287</i>	<i>+0.125</i>
Consecutive digits	<i>+0.361</i>	<i>+0.338</i>	<i>+0.188</i>	<i>+0.332</i>	<i>+0.276</i>		<i>+0.320</i>
Rearranging digits	<i>+0.349</i>	<i>+0.118</i>	<i>+0.216</i>	<i>+0.234</i>	<i>+0.158</i>	<i>+0.310</i>	

These positive intercorrelations give more definite indication of a common factor for oscillation than did the corresponding table in the previous investigation.

VII. FACTORIAL ANALYSIS OF INTERCORRELATIONS

In an effort to confirm this suggestion the above table of raw correlations was subjected to multiple factor analysis as described by Thurstone (4), appendix) and as exemplified by Alexander (1).

(1) The analysis showed that about 32% of the variance was accounted for by the first factor (see Table XII, p. 45 for loadings), and that of the residual correlations after it had been extracted, only one was as much as twice the P.E. of the original correlations. To express this in another way, when the analysis was continued to extract another factor, it accounted for only 8% of the variance, and a third factor for 7.4%.

Hence either from the point of view that the residual correlations after the first table are not significant (i.e. if the usual standards of significance hold for these residual correlations), or that any factors after the first account for a very small part of the total variance, it may be concluded that a single factor, which is not G , is sufficient to account for the observed intercorrelations.

(2) A similar result was obtained when the analysis was made with the coefficients which had been corrected for age. This time the first factor accounted for 30% of the variance (see Table XII, p. 45 for loadings), the second and third for 7 and 5% respectively. Also, only one residual correlation after the first table was as much as twice the original P.E. The elimination of age is seen to have increased the relative importance of the first factor. A more detailed account of further factor analysis will be given later.

VIII. ESTIMATION OF STEADINESS

To turn now to the other aspect of the investigation, the estimation of steadiness of character, various improvements in the methods of obtaining estimates were adopted, allowing of intercorrelations of known P.E. being obtainable.

The subjects comprized four school classes, and estimates of steadiness of character for each child were obtained from either three or four of its teachers, after at least one session's acquaintance. To assist the teachers in giving estimates of steadiness, a scale of nine points was suggested, and the number in each class was divided in a manner as nearly normal as possible among these nine positions. The teachers had then to grade the pupils relatively, putting the correct number in each position. They started from the least and most steady children, giving them estimates of 1 and 9 respectively, and so working in to the average child who would be given 5. This method overweights the extremes, but not to anything like the same extent as was previously found to be the case when the estimates were made without the guidance of a table of the relative numbers in each group. It has, moreover, the important advantages that the sets of estimates have all equal variances and the distribution of the mean estimates is not unduly narrowed ((6), p. 6, note).

The fact that all the coefficients are negative strengthens the indication of an inverse relationship given by the four significant coefficients. A single figure showing the amount of connexion between the estimates and the tests was obtained by correlating the mean σ/M scores for all the tests with the mean estimates. This gave $r = -0.320$ P.E. = 0.065 which is significant. This figure is likely to be swollen by the combined effects of age and intelligence, and it is therefore desirable to see what effect the elimination of these two variables has on the coefficient. The correlation coefficients required to calculate the required partial correlation between tests and estimates are:

(Mean Score—I.Q.) $r = +0.012$ (Mean Estimate—Age) $r = +0.154$.

(Mean Estimate—I.Q.) $r = +0.201$ (Age—I.Q.) $r = -0.279$.¹

(Mean Score—Age) $r = -0.295$.

The resulting partial correlation is:

(Score—Estimate, Age and I.Q. eliminated) $r = -0.275$,

which is still significant.

X. EFFECT OF MISTAKES

It was thought to be desirable to consider what connexion, if any, existed between the number of mistakes made by the subjects in the tests, and the other variables. In four of the tests, viz. code, rearranging digits, writing letters, writing numbers, mistakes were entirely absent, or negligible in number. In the other three they sometimes occurred to quite a large extent. Their number was seen to be roughly proportional to the speed of working, so a measure of the relative number of mistakes

was obtained by using $\frac{\text{Number of mistakes}}{\text{Mean score}}$. This measure was correlated

with the mean estimate, with the coefficient σ/M for oscillation and with I.Q., giving the coefficients shown in Table X.

Table X

Test	Correlation of mistakes with					
	Mean estimate	P.E.	σ/M	P.E.	I.Q.	P.E.
Consecutive digits	-0.164	0.07	+0.031	0.075	+0.114	0.08
Subtracting digits	+0.004	0.07	-0.101	0.07	-0.043	0.08
Adding digits	-0.099	0.08	+0.310	0.07	-0.220	0.09

¹ This negative correlation usually occurs in the school from which the subjects were taken, when ages over 14 are included; i.e. the brighter children tend to (or tended to) leave before the others. Also each class included a few older, less intelligent children repeating a year.

44 *Individual differences in amplitudes of oscillation*

Only one of these is significant, that for adding digits. This result is possibly due to the fact that adding digits was the first test to be given. This possibility is strengthened by the fact that only eight of the ninety subjects made any mistakes in the second application of this test. Hence these results cannot be taken to show the existence of any connexion between whatever is measured by mistakes and the other variables considered.

XI. TEST LOADINGS WITH THE GROUP FACTOR FOR OSCILLATION

To obtain more detailed information about the factor for oscillation common to the tests and estimates, and to discover especially its relationships with the tests, some further multiple factor analysis was undertaken. This time the intercorrelations with the tests for age, I.Q., and estimates were included, giving a ten-square matrix. From this matrix the analysis was continued until six factors, of diminishing importance, had been extracted. As previously found, the first factor was considerably more important than the others, but as was to be expected from the greater diversity of the entries in the matrix, it did not account for such a high proportion of the variance. The loadings of the matrix entries with the factors are shown in Table XI.

Table XI

	Factors						h^2	h
	I	II	III	IV	V	VI		
Estimate (mean)	-0.461	-0.192	-0.163	+0.512	+0.094	+0.014	0.547	0.740
Test: Adding	+0.582	-0.362	-0.086	+0.051	-0.215	-0.072	0.531	0.730
Consecutive digits.	+0.515	-0.216	+0.427	-0.108	+0.057	-0.109	0.523	0.725
Code	+0.654	+0.294	-0.097	+0.062	+0.153	-0.099	0.560	0.749
Writing digits	+0.559	-0.055	+0.018	-0.041	+0.218	+0.014	0.366	0.606
Writing letters	+0.658	+0.137	-0.241	+0.258	+0.127	-0.197	0.633	0.797
Subtraction	+0.480	-0.005	+0.095	-0.103	+0.053	-0.128	0.269	0.520
Rearranging digits	+0.412	-0.286	+0.107	+0.137	-0.141	+0.185	0.337	0.581
I.Q.'s	-0.020	-0.539	-0.172	-0.100	+0.479	+0.392	0.714	0.846
Age	-0.328	+0.199	+0.624	+0.109	+0.117	-0.127	0.580	0.761
Σk^2	2.504	0.732	0.723	0.402	0.411	0.285	5.060	—
$\Sigma k^2/n$	0.2504	0.0732	0.0723	0.0402	0.0411	0.0285	0.506	—
%	25.04	7.32	7.23	4.02	4.11	2.85	50.6	—

These loadings do not give any indication that any one test, or the mean estimate, is identical with the first common factor, or with any of the others. Hence there was no advantage to be gained by rotating the axes to coincide with any particular test or tests. Instead the first factor loadings were compared with the correlations between the entries in the table and the mean σ/M for oscillation, giving Table XII. The last row of this table gives the loadings in an analysis of the seven tests and mental age, and shows that the effect of mental age is similar to that of

chronological age, as was to be expected from the intercorrelations already given (Tables V and VI).

Table XII

	Test							Mean estimate	Age	I.Q.
	1	2	3	4	5	6	7			
Correlation with mean σ/M	+0.70	+0.60	+0.62	+0.52	+0.67	+0.61	+0.58	-0.32	-0.295	+0.01
First factor loadings from Table XI	+0.58	+0.48	+0.66	+0.56	+0.65	+0.52	+0.41	-0.46	-0.33	-0.02
First factor loadings from § VII (1)	+0.58	+0.49	+0.65	+0.58	+0.62	+0.55	+0.44	—	—	—
First factor loadings from § VII (2)	+0.55	+0.49	+0.61	+0.55	+0.59	+0.60	+0.41	—	—	—
First factor loadings (7 tests and mental age)	+0.60	+0.47	+0.67	+0.55	+0.67	+0.50	+0.42	—	Mental age -0.36	

The third and fourth rows give the corresponding loadings of each test with the first common factor from the analyses mentioned in § VII. It is seen from the above table that the loadings of the tests, from the four analyses, are in good agreement with one another, and that they, together with the loadings for age, I.Q. and estimates, are also in fair correspondence with the mean σ/M correlations. This correspondence gives further indication that the first, and most important, factor common to the tests is that of oscillation. To verify this point in yet another way, a further analysis was made, this time of a table containing the intercorrelations of the mean scores for oscillation (σ/M), the seven tests, and the mean estimates. The first common factor loadings and the correlations with the mean score for oscillation are shown in Table XIII.

Table XIII

	Mean score	Test							Mean estimate
		1	2	3	4	5	6	7	
Loadings of first common factor	+0.97	+0.62	+0.61	+0.645	+0.57	+0.63	+0.55	+0.48	-0.39
Correlation with mean σ/M	—	+0.70	+0.60	+0.62	+0.52	+0.67	+0.61	+0.58	-0.32

In this analysis the first common factor accounted for 39% of the variance, and the next for 6%. The loadings are nearer the correlations with the mean σ/M than in the other analyses, and the high loading of 0.97 of the mean score points again to the common factor present in the tests being that of amplitude of oscillation.

XII. SUMMARY

It has been shown that:

(1) A common factor other than G is present in the tests given and the estimates obtained for steadiness of character, and that this group factor is most closely identified with the mean score for amplitude of oscillation.

(2) This group factor is significantly and negatively correlated with estimates of steadiness of character.

(3) This group factor is independent of I.Q., and also is probably independent of what is measured by mistakes.

(4) For the age range of the subjects used ($11\frac{1}{2}$ – $16\frac{1}{2}$ years) this group factor shows a significant decrease with age and also with mental age. Its presence is still evident when the effect of age has been eliminated from the intercorrelations.

In conclusion, the writer wishes to express his thanks to Dr R. H. Thouless for his assistance in the discussion and analysis of the results.

REFERENCES

- (1) ALEXANDER, W. P. (1935). "Intelligence, concrete and abstract." *Brit. J. Psychol. Monog. Suppl.* VI, No. 19.
- (2) FISHER, R. A. (1936). *Statistical Methods for Research Workers*. Edinburgh: Oliver and Boyd.
- (3) PUROSHOTTAM (1932). "Individual differences in fluctuations of attention." London University, Ph.D. Thesis (unpublished).
- (4) THURSTONE, L. L. (1935). *The Vectors of Mind*. Chicago.
- (5) WALTON, R. D. (1936). "Oscillation and steadiness of character." *Brit. J. Psychol.* XXVII, 181.
- (6) VERNON, P. E. (1938). "The assessment of psychological qualities by verbal methods." *Industrial Health Research Board Rep.* No. 83. London: H.M. Stationery Office.

(Manuscript received 6 June 1938)

CHANGES IN FIGURE-GROUND PERCEPTION IN PATIENTS WITH CORTICAL LESIONS

BY M. R. HARROWER

From the Department of Neurology and Neurosurgery, McGill University and Montreal Neurological Institute.

- I. *Introduction* (p. 47).
- II. *Method of investigation* (pp. 47-48).
- III. *Results* (pp. 48-50).
- IV. *Discussion* (pp. 50-51).
- References* (p. 51).

I. INTRODUCTION

IN a previous investigation⁽³⁾ an attempt was made to isolate the determining factors in figure-ground articulation using modifications of the Rubin profile-vase drawing⁽⁴⁾. Perception of one part of the field as figure seemed to depend, with normal subjects, on its smallness, surroundedness, the good continuation and closure of its lines, its greater articulation. It was also found that a truly ambiguous figure, i.e. one in which none of these factors was operating, would be seen in terms of a previous perception in which one part of the field had definite figural properties; but this influence from a previous perception, established attitude or 'carry-over' from one figure to another, did not occur when the second drawing possessed other and opposite figural properties. In other words, if an 'ambiguous' figure followed one in which the profiles had been enhanced the tendency was for it to be seen as profiles also, but even a slight enhancement of the vase was sufficient to destroy any effect of the previous perception.

II. METHOD OF INVESTIGATION

The present investigation deals with the question of figure-ground perception in patients with cerebral lesions. The performance of 30 patients with cerebral lesions (the experimental group) was compared with 30 'normal' patients, i.e. patients without brain involvement, from the same hospitals. The experiments were carried out at the Montefiore Hospital in New York, and the Neurological Institute in Montreal.¹

¹ The author wishes to express thanks to Dr K. Goldstein for his interest and encouragement in this study; and to Drs W. Penfield and W. V. Cone for the co-operation which made it possible.

No differentiation on the basis of localization of the lesion was made in this investigation. The patients were in the main post-operative cases with unilateral lesions, both hemispheres, frontal, frontal parietal, and temporal lobe cases being represented.

Fig. 1¹ shows the modifications of the original drawing which were used. These consisted of a series of seven in which three enhanced the profiles, three the vase, while one was an ambiguous figure in which profile or vase could be seen equally easily by normal subjects. These drawings, on large cards, were presented singly to the patients. The order of presentation was either:

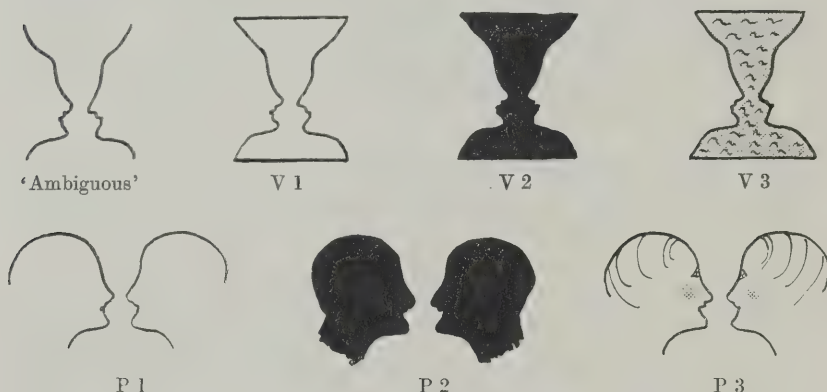


Fig. 1.

Ambiguous, Profile 1, P 2, P 3, Ambiguous, Vase 1, V 2, V 3, Ambiguous; or, Ambiguous, Vase 1, V 2, V 3, Ambiguous, Profile 1, P 2, P 3, Ambiguous. When the first Ambiguous figure was seen as a Vase, the subsequent card was Profile 1, and vice versa.

The instructions were simply that the patient should tell what he saw.

III. RESULTS

A normal record was considered to be one in which the ambiguous figure was seen either as profile or vase, the subsequent drawings in terms of their figural enhancements, the repeated ambiguous figures either in terms of how they were first seen, or modified by an attitude created by the preceding drawings. Twenty-seven, or 90%, of our normal patients gave such records; only two patients, or 6.6%, in the experimental group gave 'normal' records.

¹ V 3 was red in the original, with black markings; P 3 had red cheeks, and blue eye markings.

This leaves the twenty-eight records in the experimental group (and the three who deviated in the normal group) to be considered to see what light they can throw on figure-ground perception in cases of cerebral lesions.

The deviations from the normal record were of four kinds:

I. There is a marked perseveration of the attitude engendered by the first perception, such that the figure seen in the first perception is maintained despite counter influences of figural properties.

Example. If the first ambiguous figure is seen as profiles, no vase will be seen in the subsequent drawings despite their 'vase enhancement'.

II. There is a failure to recognize the ambiguous figure as any meaningful object, even after correctly perceiving all the other figures.

III. There is a rigidity such that the less dominant figure cannot be seen at will, despite its having been perceived in other drawings in which it was the dominant figure.

IV. There is an introduction of other objects in the place of profiles or vase, either throughout the whole series or in parts of it.

Example. One subject in the experimental group saw such a variety of objects as 'a mound of hay' (Amb.) (P. 1); 'Priests' (P. 2); 'People' (P. 3); 'a table' (V. 1); 'a priest's hat' (V. 2); 'a cardinal's hat' (V. 3).

Table I gives the distribution of these deviations among the experimental and control groups.

Table I

	Cortical group	Controls
Perseveration:		
3 places	4	0
2 places	4	0
1 place	2	2
Total	10	2
Failure:		
On ambiguous figure	8	1
On other figures	4	0
On all except P. 3	1	0
Total	13	1
Introduction of new objects	6	0
Inability to shift	4	0
Normal records	2	27
Total no. of deviations*	33	3

* Since some records deviated from the normal in more than one way, the total of such deviations in the cortical group is 33; i.e. 28 patients out of the 30 showed one or more deviations.

Table II shows examples of typical deviations as they appear on the records.

Table II

Order of presentation	Perseveration	New objects	New objects	Failures	
Amb.	P P P		Fly	-	-
V. 1	P P P		Mushroom	V	-
V. 2	P P V		Tulip	V	-
V. 3	P - * V		V	V	-
Amb.	P P P V	Hay	-	-	-
P. 1	P P P V	Hay	Octopus	P	P -
P. 2	P P P P	Priests	P	P	P -
P. 3	P P P V	People	P	P	P P
Amb.	P P P V	Hay	-	-	-
V. 1		Table		V	
V. 2		Priest's hat		V	
V. 3		Cardinal's hat		V	
Amb.		Hay		-	

* - stands for failure to perceive any object.

Two other characteristics of the responses of the cerebral group may be mentioned here. Whereas the normal subjects frequently mentioned casually that these drawings were such that both figures could be seen in quick succession, even if only one 'hit you in the eye' to start with, only one patient in the cerebral group ever made such a comment, and in this case it was for him a genuine discovery which caused him considerable excitement. This does not mean that these patients failed to have both experiences, but the dual aspect, the ambiguity, the variation on a central theme, seems to have been less striking to them. The other characteristic of the records of the cerebral group was the more frequent reference to concrete details in the pictures; the 'man and woman', Roosevelt, Washington, the red cheeks and black hair of Profile 3, etc., were commented on.

IV. DISCUSSION

Brief mention must be made of the theoretical implications of these results. We have found that in these patients a simple perceptual experience has different properties from those found in normal subjects. As many as 43% are unable to make anything out of the ambiguous figure, a drawing, however, which presents no puzzle for normals. One third of the patients show the influence of an established attitude, a tenacious holding to the impression created at first, in the face of a counteracting perceptual influence, and an inability to pick out the two aspects of the figure simultaneously. We also find the rather surprising introduction of other, and not necessarily relevant objects, and the lack of ability to see the series as a series. But what do these things mean?

We must assume, I think, from the failure on the ambiguous figure that spontaneous sensory organization differs in such patients in that a

given retinal stimulus distribution, which would produce a well organized experience in normals, may result in chaos, failure, for some of these patients.

Secondly, the introduction of familiar objects, as for example the mound of hay and the priests seen by the catholic farmer, the lack of ability to see the series in its serial character, the emphasis on concrete detail, all point to the fact that these patients are less capable of taking a detached attitude in regard to their perceptions, and, where possible, incorporate them in a framework of tangible and concrete reality... evidencing perhaps what Goldstein^(1,2) has called the 'loss of abstract attitude'.

Finally the marked perseveration of the original perception, the clinging to the established attitude, seems to be also a matter of this lack of detachment... the Ego, the core of the individual one might say, is involved to a greater degree even in such simple performances, and 'failure' becomes a dangerous situation to be avoided, while the perpetuating of a successful response becomes a safeguard against possible difficulties. This appears most strikingly in certain individual patients who show real distress on being unable to make anything out of the ambiguous figure. Again, those patients who only see one figure through-out frequently assert that response with undue vehemence.

The value of such findings would seem to lie in the fact that they may give us clues to certain qualitative differences in the experiences and behaviour of persons suffering from cerebral lesions, differences which might well leave unaltered a quantitative estimate of the intelligence level, and might well pass unnoticed in ordinary daily life, and yet if recognized may enable us to reconstruct more accurately the changed psychological world resulting from changes in the brain.

REFERENCES

- (1) GELB, A. & GOLDSTEIN, K. (1920). *Psychologische Analysen hirnpathologischer Fälle*. Leipzig.
- (2) GOLDSTEIN, K. (1936). "The significance of the frontal lobes for mental performances." *J. Neurol. Psychopath.* xvii, 27-40.
- (3) HARROWER, M. R. (1936). "Some factors determining figure-ground articulation." *Brit. J. Psychol.* xxvi, 407-24.
- (4) RUBIN, E. (1921). *Visuell wahrgenommenen Figuren*. Copenhagen.

(Manuscript received 21 July 1938)

BINAURAL LOCALIZATION WITH TWO NOTES DIFFERING IN PHASE BY 180°

By J. W. HUGHES

From the Physiology Institute, Newport Road, Cardiff

- I. *Preliminary experiments* (pp. 52-53).
- II. *Results of main experiments* (pp. 53-56).
- III. *Previous work* (p. 56).
- IV. *Summary* (p. 56).

I. PRELIMINARY EXPERIMENTS

It is well known that if two notes of the same pitch are presented to the two ears, and it is arranged that they are of equal loudness and are in phase with one another, the sound which is heard appears to come from a source located somewhere in the median plane, though different observers vary as to its position in that plane. During some experiments in which the phase of one of these notes was changed by 180° , it was found that the sensation then produced by the two notes was quite different from that in the first case. The observer was conscious of a stimulus applied to each ear; the sensation was symmetrical but the source was no longer localized sharply in the median plane. When a sudden change from zero phase difference to 180° phase difference was made, a broad, diffused source of sound seemed to have been substituted for the apparent point source of the first case. This change was easily detected by the observer. The two telephones employed were fed from two separate secondaries on the output transformer of the amplifier following the heterodyne oscillator from which the pure tones used were obtained, and the connexions to one of these secondaries could be reversed by a switch. Interruptions of the telephone currents were made for about 1 sec. every 5 sec. During any one of these intervals of silence, a reversal could be made unknown to the observer; such changes were very rarely missed. The observer sat in a soundproof room and signalled when he considered that a change had been made. The frequencies used were 200, 400, and 3000 c./sec., and tests were made on three observers.

An investigation of the change of intensity required in one ear to produce a definite localization to one side or the other was next made for

both cases, in the expectation that one might prove more sensitive to such changes than the other.

Two amplifiers supplied the telephones, which were each in series with a 1-megohm resistance. One of these circuits was connected to a centre-tapped transformer, with the centre-tap earthed, and a switch for connexion to one or other of the outer terminals for phase-changing; the terminal not in use was earthed through another 1-megohm resistance. The voltages applied to the output circuits were arranged to be in phase by using a cathode-ray oscillograph in conjunction with a variable condenser ($0.0015\ \mu\text{F. max.}$) across one of the circuits. Since the output circuits are of high resistance and low inductance, the currents in them, and therefore in the telephones, will differ very little in phase.

In the first results taken, the intensities used were sufficient to give a moderately loud note, say 30–40 db. above threshold. After the phase correction had been made and the intensities approximately equalized by obtaining a straight line at 45° to the axes on the oscillograph, small intensity adjustments were made to obtain 'central' localization. Then a series of readings was taken; the intensity was reduced slowly in each ear in turn, and the observer signalled when he detected a change in the localization of the combined note. The oscillograph showed that there was no phase change with the intensity changes actually needed. In every case the observer noticed the 'return swing' of the combined note when the original symmetrical conditions were restored. Results were taken at 400 and 800 c./sec. in this way with two observers.

In the ordinary case of zero phase-difference, change of intensity in one ear sufficient to disturb the central localization causes an apparent angular shift of the sound out of the median plane. In the present case no such steady angular shift occurred; the change was from symmetry to a note located definitely in one ear.

II. RESULTS OF MAIN EXPERIMENTS

A further series of results was then taken with known intensities of the notes presented to the ears. Shunts were used on the telephones, and were adjusted to give monaural threshold intensity at 0 db. on the attenuator in each amplifier. Then settings were made as before with the attenuator at +50, +40, +30 and +20 db. It was found that it was better to make the intensity adjustment before starting when the phases differed by 180° . Ten settings were made for each result in Table I, and the same two frequencies were employed. Most of the results were taken by two observers; one of them (J. W. H.) required far greater changes than the other (W. L. B.) though the 'return-swing' was easily detected by the latter. In fact, he complained that the rate of decrease of intensity

was too rapid in the first trials, and it was reduced when taking the results shown. Some settings by a third observer (J. H. S.) agreed roughly with those of the second.

Table I. *Gradual intensity change (db.) for just perceptible change of location*

Observer ...	J. W. H.				W. L. B.				J. H. S.			
Frequency ... (c./sec.)	400		800		400		800		400		800	
Intensity	0°	180°	0°	180°	0°	180°	0°	180°	0°	180°	0°	180°
+50 db.	—	—	4.3	3.1	—	—	—	—	—	—	—	—
+40 db.	2.0	2.5	4.8	2.6	1.1	0.9	0.8	1.0	0.5	0.4	2.4	1.4
+30 db.	5.8	2.6	5.4	3.0	1.1	0.8	1.1	1.1	—	—	—	—
+20 db.	7.5	4.5	6.7	3.9	1.1	1.0	0.9	1.0	—	—	—	—
+30-40 db.	7.8	5.3	11.8	3.9	1.0	1.3	1.0	0.7	0.7	0.7	1.4	1.7
	—	—	2.2	1.1	—	—	0.7	0.5	—	—	—	—
	5.8	4.5	2.1	1.2	1.2	1.2	0.8	0.7	—	—	—	—

To find whether this discrepancy was due to some difference between the observers' uniaural differential thresholds, these were determined for each observer at the frequencies and intensities employed in the foregoing. The output of one amplifier was used, and a fixed shunt of $1000\ \omega$, with a variable resistance R in series with it, was used on the telephone. It was arranged that R was short-circuited for about half a second every second, so that the intensity in the telephone rose and fell twice a second. The note was presented to the observer at the required intensity, and the value of R was reduced or increased in accordance with his signals to determine the value of the uniaural differential threshold. Then this threshold is given approximately in decibels by $20 \log_{10} (1 + R/1000)$. Five settings were made at each intensity; if a rest were required, the observer could silence the telephone between settings by a switch. It will be seen from Table II that there is little difference between the results for the three observers.

Table II. *Uniaural differential thresholds (db.)*

Observer ...	J.W.H.		W.L.B.		J.H.S.	
Frequency (c./sec.)...	400	800	400	800	400	800
Intensity						
+50 db.	0.7	0.6	0.6	0.4	0.4	0.6
+40 db.	0.9	0.7	0.7	0.4	—	—
+30 db.	1.0	0.5	0.7	0.4	—	—
+20 db.	0.9	0.7	0.7	0.4	0.4	0.5

Using this latter form of the apparatus, the detection of changes of location caused by sudden change in intensity was studied. The notes were balanced in intensity as in the first experiments, and then R was

increased from zero to about 100ω . The location of the note then changed from the median plane to one side when the phase difference was zero, or from the broad and diffused type of source to a note definitely in one ear when the phase difference was 180° . As R was reduced according to the observer's requirements, this change gradually disappeared, and the value of R when it could no longer be detected was found. The variations in intensity caused by R were still noticeable at this stage. In these experiments it was found advisable to double the times for which R was open- and short-circuited. The results for two observers at various intensity levels and two different frequencies are given in Table III. The intensity changes needed by the second observer (W. L. B.) are again lower than those required by the first (J. W. H.), and though they are of the same order as his differential thresholds, he (W. L. B.) maintained that intensity variations were still audible at the settings made.

Table III. *Sudden intensity change (db.) for just perceptible change of location*

Observer	J. W. H.				W. L. B.			
Frequency (c./sec.) ...	400		800		800		400	
	0°	180°	0°	180°	0°	180°	0°	180°
Intensity								
+50 db.	4.0	2.5	2.3	2.3	0.6	0.5	—	—
+40 db.	2.8	2.0	2.3	2.8	0.6	0.6	0.6	0.6
+30 db.	3.2	4.3	2.9	2.3	0.7	0.6	1.1	0.9
+20 db.	3.5	2.9	2.6	3.2	1.0	0.8	0.9	0.8

A comparison between the just perceptible gradual changes of phase for differences of phase between the notes of 0° and 180° was also made. These tests were carried out during some work on the variation of the amount of these changes with frequency; it is hoped that a fuller account of the apparatus and the methods used will be published shortly. Both phase and intensity of either note could be independently varied. The phases of the notes were equalized (or opposed) before the start of an experiment, and the intensity of one varied a little to give 'central' localization as already described. At a given signal, the phase of one note was slowly changed until the observer signalled that a change of location had been detected. Precautions were taken to ensure that no external criteria were used. All the measurements were taken with notes at 50 db. above threshold.

The results for the three observers indicate that there is no significant difference between the least perceptible changes of phase detectable with either phase relation at a given frequency. Both show an increase in the amount of change required as the frequency is increased, and at about

1500 c./sec. it becomes impossible to detect a change of phase with either phase relation. A typical set of figures is shown in Table IV; they seem to indicate that failure occurs at a rather lower frequency when the notes differ in phase by 180° . These deductions are confirmed by some further experiments in which sudden phase changes were made and the least perceptible change was found.

Table IV. *Gradual phase change (degrees) for just perceptible change of location (J. W. H.)*

Frequency (c./sec.)	...	600	700	800	900	1000	1100	1200	1300	1400
Phase difference	0°	24.5	28.0	23.3	23.7	20.8	22.2	31.0	59.0	—
	180°	30.0	20.0	28.3	27.5	21.6	30.0	108.0	113.0	—

III. PREVIOUS WORK

It is of interest to note that in the experiments of Halverson¹ on binaural localization the effect of varying the relative phases of the two notes was investigated, and the subjective phenomena when the difference was 180° were also noted. His observers stated that the sensation was one of a double sound of equal intensity at the two ears, and that there was no clear image in the median plane. As the phase was varied slowly through the position of 180° phase difference, the combined sound did not pass through the head; the observer was conscious of two distinct images on the aural axis, and as the phase changed one approached the median plane from one side, while the other receded and finally disappeared. Though the observers in the present work were unaware of these experiments and made no special introspections as to the nature of the change, the descriptions of the effect offered by them and given above are much the same as those of the trained introspective observers of Halverson.

IV. SUMMARY

The least perceptible variations in the relative phases and intensities of two notes differing in phase by 180° and 0° have been compared, and little difference found between them. The sensations experienced in the first case are described.

The author wishes to thank the Medical Research Council for grants which made the work possible, and to express his appreciation of the suggestions made by Dr J. H. Shaxby, in whose laboratory the experiments were performed.

¹ Halverson, H. M. (1922). *Amer. J. Psychol.* xxxiii, 178.

SOME RACIAL DIFFERENCES IN PERCEPTION

By W. M. BEVERIDGE

- I. *Phenomenal regression for brightness and whiteness* (pp. 57-59).
 - (1) *Method of experiment* (pp. 57-58).
 - (2) *Results* (pp. 58-59).
- II. *A picture preference test* (pp. 59-60).
- III. *Mental compromise in estimating direction* (pp. 60-64).
 - (1) *Method of experiment* (pp. 60-62).
 - (2) *Results* (pp. 62-63).
 - (3) *Relationship of mental compromise in the estimation of direction with other factors* (pp. 63-64).
- IV. *Summary* (p. 64).
References (p. 64).

I. PHENOMENAL REGRESSION FOR BRIGHTNESS AND WHITENESS

(1) *Method of experiment*

In a former paper⁽¹⁾ it was shown that West African students show a greater tendency to phenomenal regression for shape and size than do British students. The aim of the following experiments was to find out if these African students showed also a greater tendency to phenomenal regression in the perception of brightness and whiteness.

As in the earlier tests the subjects were natives of the Gold Coast, Ashanti and Togoland, male students of the Presbyterian Training College, Akropong; on the average they were about eighteen years of age. The procedure was a copy of that used by Dr Thouless in his experiments with Scottish students⁽²⁾. A white disk 4 in. square was supported at right angles to the table and to the subject's line of vision, at a distance of 290 cm. from the source of light. A grey disk of the same size was arranged so that the subject saw it to one side of the white disk and separated from it. The distance of the grey disk from the lamp was adjusted until the subject reported that the two disks looked equal in brightness or whiteness as the case might be.

The disks were cut from Zimmermann's series of fifty grey papers, one being the standard white and the other no. 13 grey. The latter had a reflectivity 0.228 that of the white. The sole source of light was a $4\frac{1}{2}$ v. electric bulb, and the subject sat behind the lamp.

The instructions were: "That is a white disk, and this is a grey one. When I hold it here (holding the grey disk well behind the point at which its luminosity equalled that of the other) it looks less bright than the other. As I bring it nearer to the lamp, it looks brighter and brighter, until finally it looks brighter than the other one. Now I want to find the exact position at which it looks the same brightness as the other one."

After preliminary practice three ascents and three descents were made, and the mean was taken to represent the position at which the disks seemed equally bright for that subject. The same procedure was used to determine the position of equal whiteness. On each ascent and descent the subject balanced both for brightness and for whiteness. As in the earlier tests the measure used was the 'Index of phenomenal regression' (P.R.), i.e. $(\log P - \log S)/(\log R - \log S)$, where P is the numerical value of the phenomenal character, S that of the stimulus character, and R that of the real character.

(2) Results

Forty subjects were tested. None of them failed to make a considerable difference between the position for brightness and that for whiteness. Only in one case was the ratio of P.R. for whiteness to that for brightness less than 1.35. This makes an interesting contrast to Dr Thouless's results(2); out of forty-five Scottish students he found twenty-two for whom that ratio did not exceed 1.35, and for ten of these the position of equal phenomenal brightness was also that of equal phenomenal whiteness.

The mean African P.R. for brightness was found to be 0.281. The European mean found by Dr Thouless was 0.325. The difference is 0.044, and the P.E. of difference 0.015. There was only one negative case. Thus, contrary to expectation, the African students showed *less* P.R. for brightness than did the European students. The medians and quartiles are shown in Table I.

Table I

Group	No.	Lowest	First quartile	Median	Last quartile	Highest
Index of P.R. for brightness						
European	104	(Negative)	0.19	0.305	0.47	0.66
African	40	(Negative)	0.15	0.28	0.39	0.56
Index of P.R. for whiteness						
African	40	0.548	0.666	0.708	0.729	1.00

The mean African P.R. for whiteness was 0.698. One subject saw the grey paper still grey even when it was touching the lamp. Dr Thouless publishes a list of only eight subjects' P.R. for whiteness; their mean is 0.606. This is 0.092 less than the African. As the sample is so small we

may apply the χ^2 test, forming a fourfold table about the median of the group as a whole(3). With Yates's correction for continuity we find χ^2 equal to 3.75. The likelihood of this occurring by chance sampling is practically one in twenty.

Thus we find that while the African students have less P.R. for brightness than the Scottish students, they have probably more for whiteness. It was whiteness, not brightness, however, which Dr Thouless found to correlate highly with phenomenal regression for size and shape. It is interesting to note that with the Africans there seems to be a small *negative* correlation (-0.27 ± 0.098) between P.R. for brightness and P.R. for whiteness.

II. A PICTURE PREFERENCE TEST

The aim of this experiment was to discover if the African's greater tendency to phenomenal regression affected his appreciation of pictorial art. Dr Thouless has shown that certain features of Oriental art, "the absence of shadows, and the partial or total absence of perspective", are probably due to the fact that Orientals have a large tendency to phenomenal regression(4). These features are seen also in the work of African artists such as Enwonwu, Ibrahim, Nnachy, and Ibeto(5,6), and in the drawings of African school children and college students. Do Africans, therefore, regard these features as merits? Do they prefer Oriental pictures to European?

The subjects were again Gold Coast students. The material used was a set of eighty coloured postcards, forty of famous European pictures, and forty of Japanese, Indian, and Persian pictures. (The former were mostly published by The Medici Society, and the latter by the trustees of The British Museum.) These pictures were paired roughly according to subject, and exhibited on a board, eight at a time, the European picture being sometimes on the right hand and sometimes on the left of the Oriental one, e.g.

Sheltering from a Shower (Kiyonaga)	The Little Street in Delft (Vermeer)
From the May Border (Whittington)	Bird on a Rose-tree Branch (Hokusai)
Child Upsetting a Goldfish Bowl (Utamaro)	Pas Mèche (Bastien-Lepage)
Head of a Girl (Vermeer)	Two Girls (Shigemasa)

The subjects were unaware of the fact that one picture in each pair was European and the other Oriental.

Each student was asked to state on a slip of paper which picture in each pair he preferred and also which one of the eight he liked best of all. About a hundred men took part, but not all voted on every set of pictures; 2827 votes were cast.

The results were contrary to expectation. Western pictures received 2480 votes (86.5% of the total number), while Oriental pictures received only 337 votes (13.5%). Only two Oriental pictures were more popular than the corresponding European ones; "Ten Thousand Acres" by Hiroshige was preferred to Farquharson's "Dawn", and the "Persian Miniature from the Fable of Kalilah wa Dimnah" was more popular than Margaret Tarrant's "Sylvan Melody". In no case was an Oriental picture the favourite on the board.

These results show clearly that, in spite of high phenomenal regression, African students prefer European to Oriental art. We must remember, however, that these subjects had been used to seeing European pictures from their earliest school-days, whereas Oriental art was new to them. Had they been untouched by European influences the results might have been quite different.

It is interesting to note some of the reasons given for liking pictures. Of these the most important seemed to be the illusion of reality: "Anybody looking at the picture thinks that parts of the paper are bulging out"; "The work does not look flat", "The picture represents things naturally so as to make people forget themselves when they are looking at it." A glance at the most popular picture of each set confirms this: "From the May Border" (Whittington), "The Angelus" (Millet), "La Laitière" (Greuze), "Flatford Mill" (Constable), "The Cornfield" (Constable), "The Boyhood of Raleigh" (Millais), "Dutch Interior with Soldiers" (de Hooch), "Opal Morning" (Mary Holden Bird), "The Duchess of Milan" (Holbein), "The Young Warrior" (Rembrandt).

Admiration for the artist's technique was also frequently expressed: "The quality and folds of the cloth are very well shown"; "The colours merge very gradually"; "The light and shade are very skilfully painted"; "One can see the pattern on the lady's dress and the dazzling earring". Some pictures were liked because they told a story, and others because they seemed to represent ideal conditions: "Everything in the room is neat and in its proper place"; "The lady's hair is beautifully brushed and combed"; "The building is a masterpiece of modern architecture; it is like the palace of a great monarch". This last was a reference to Vermeer's "Little Street in Delft"!

III. MENTAL COMPROMISE IN ESTIMATING DIRECTION

(1) *Method of experiment*

The following test was devised to find how far one's judgment of the horizontal depends on visual clues; to study individual and racial

differences in this; and to discover if it is connected in any way with age, intelligence and temperament.

The apparatus used was a large wooden cupboard, 6 ft. 6 in. by 54 by 20 in. It was mounted so that it could be easily tilted to an angle of 25° . A stick, $29\frac{1}{2}$ by $\frac{3}{4}$ in. by $\frac{3}{4}$ in., was hinged to the end of the cupboard 29 in. from the floor so that it could swing up and down; this was connected with a pointer which moved across a dial on the exterior of the cupboard. There was a small translucent window 9 in. square in the roof. The interior was dark, unpainted wood.

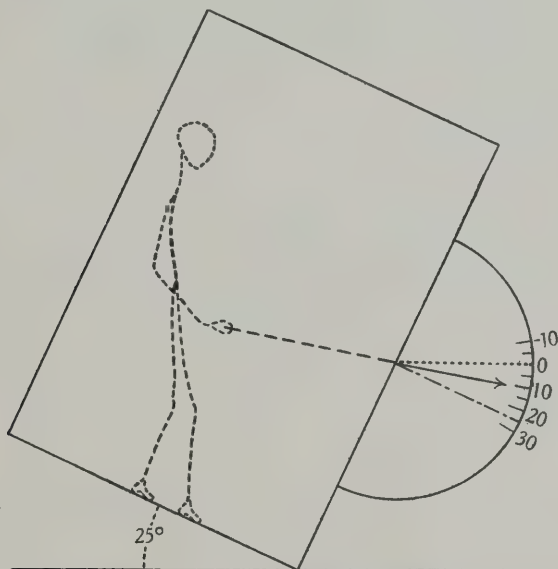


Fig. 1. On the indicator shows the horizontal and — — — — — the direction of the floor of the cupboard.

The subject entered the cupboard when it was upright, was shown the stick and its movement, and given the following instructions:

When the door is shut the cupboard will be slowly tilted; hold on to the sides while this is being done; then take the stick and hold it at what *seems* to you to be level with the ground outside, so that if a dish full of water were sitting on the stick, the water would not run out. Do not try to calculate, but hold it at what *looks* level to you.

After two or three preliminary experiments the subject raised the stick and brought it down to apparent level, then lowered it and brought it up to apparent level. Five ascents and five descents were made in this way, and the mean was taken to represent the phenomenal level for that subject.

The subjects were fifty-two West African male students, on the average about 19 years of age, and twenty Europeans, ten men and ten women, ranging in age from 25 to 50 years, the average being about 35.

(2) *Results*

With six exceptions (two Europeans and four Africans) there was a large mental compromise; the visual clues afforded by the walls, floor and ceiling of the cupboard greatly affected the judgement. As a simple measure of this let us take the error of the subject's estimation divided by the angle of tilt of the cupboard, 25° . It is obvious that if a subject were guided entirely by the visual clues he would hold the stick at an angle of 25° off level, and his score would be 1.00. If he were unaffected by visual clues he would hold the stick level and his score would be 0.00. Apart from the six exceptions mentioned above, whose scores were negative, all the scores lay between these extremes. The European mean was 0.498 (12.45° of error); the African mean was 0.418 (10.45° of error). There were great individual differences, the results lying roughly on the normal curve. The medians and quartiles are shown in Table II.

Table II

Group	No.	Lowest	First quartile	Median	Last quartile	Highest
African	52	(Negative)	0.219	0.433	0.619	0.900
European	20	(Negative)	0.434	0.543	0.642	0.858

The subjects whose scores were negative present an interesting problem. They may be compared with those whose scores were negative in the 'Brightness' test for phenomenal regression(2). In this present case, however, there must presumably be some kind of overdone reaction against what is felt to be the misleading nature of the visual stimuli.

In most cases the subject's judgement in this experiment was not so stable as in those for phenomenal regression. Complicating factors such as mental set, and possibly fatigue and health, may affect the judgement. In actually performing the experiment one has a feeling of being slightly dazed or bewildered, probably because sets of perceptual cues, usually found to be in agreement, are here in conflict.

Considering these facts we must be cautious in drawing conclusions from this test. As Table II shows, however, there seems to be a slight racial difference, the Africans on the average estimating level more accurately than the Europeans. This was to be expected; their greater tendency to phenomenal regression suggests that their perception is guided less by visual and more by other cues than is that of the European.

(In teaching and in giving intelligence tests to Africans I have found them apparently less interested in and efficient with pictures and diagrams than most Europeans, and proportionately more so with words.)

To test the significance of this apparent racial difference we may form a fourfold table round the median of the two groups taken as a whole (see Table III).

Table III

	No. below median	No. above median
Africans	30	22
Europeans	6	14
$\chi^2 = 4.43$		

A value of χ^2 is thus obtained which would not occur by the chances of sampling from identical populations more than once in about twenty-seven times. This apparent racial difference may be due in part to the European subjects being older; but, on the other hand, they were almost certainly also more intelligent, and these two factors would tend to cancel each other out. We must, however, remember the probable effect of practice; from early childhood Africans are used to carrying loads balanced on their heads.

(3) *Relationship of mental compromise in the estimation
of direction with other factors*

(a) *Age.* The African subjects varied little in age. With the Europeans there seemed to be a tendency for age to increase the ratio of mental compromise. The correlation was 0.517 ± 0.112 . From Fisher's Table V_A(3) we find a value of P practically equal to 0.02. Taking into consideration the similar connexion between age and phenomenal regression, we have the suggestion that the effect of experience on perception increases with age at the expense of the effect of the stimuli actually present.

(b) *Intelligence.* On the basis of personal knowledge and their College record, 24 of the Africans were ranged in order of intelligence. A correlation of -0.41 was found between intelligence and ratio of mental compromise in estimating direction. This is just significant ($P = 0.05$).

(c) *Temperament.* Three people working independently arranged the twenty Europeans into five classes ranging from very extroverted to very introverted. As the correlation between the three gradings was high, the combined grading may be regarded as fairly accurate. There seems to be a connexion between extrovert tendencies and mental compromise in direction. The correlation was 0.486 ± 0.114 . From Fisher's Table V_A we

find that P is roughly 0.03. As a further test, fourteen Africans were chosen by two observers as being extroverted and fourteen as being introverted. A fourfold table was formed giving a value of $\chi^2=2.33$. This would occur in chance sampling about once in seven times, but it tends to confirm the previous result.

(d) *Phenomenal regression*. There seems to be a positive connexion between this and tendency to mental compromise in direction. Unfortunately only twenty subjects did both tests, but a value of P was found equal to 0.036.

IV. SUMMARY

1. African students show less phenomenal regression for brightness, more for whiteness, and therefore more discrimination between the two, than do British students.

2. This great tendency to phenomenal regression is reflected in African art, but it does not lead African students to prefer Oriental to European pictures.

3. Our perception of the horizontal depends to a considerable extent upon visual clues. When these contradict the more reliable internal clues, most people make an unconscious mental compromise.

4. There are very large individual differences in the amount of this mental compromise.

5. There is very probably also a racial difference, Africans being less affected by the visual clues than Europeans.

6. Mental compromise in estimating direction has a positive correlation with age, with extrovert tendencies of temperament, and very probably with phenomenal regression. It probably has a negative one with intelligence. Further investigation is, however, necessary.

REFERENCES

- (1) BEVERIDGE, W. M. (1935). "Racial differences in phenomenal regression." *Brit. J. Psychol.* xxvi, 59.
- (2) THOULESS, R. H. (1932). "Individual differences in phenomenal regression." *Brit. J. Psychol.* xxii, 216.
- (3) FISHER, R. A. (1936). *Statistical Methods for Research Workers*.
- (4) THOULESS, R. H. (1933). "A racial difference in perception." *J. Soc. Psychol.* iv, 330.
- (5) *Nigeria*, June 1938 (illustrations).
- (6) *The Listener*, 21 July 1937 (illustrations).

(*Manuscript received 28 October 1938*)

PUBLICATIONS RECENTLY RECEIVED

General Psychology from the Personalistic Standpoint. By WILLIAM STERN. Translated by H. D. SPOERL. New York: Macmillan. 1938. Pp. xxii+589. 20s net.

Titchener has often been held out to us as the Englishman who represented the German psychological tradition in America, but it is seldom recognized that the late William Stern has a strong claim to the title of the Jew who represented the British psychological tradition in Germany. If only for this reason, Dr Spoerl's distinguished translation of the *Allgemeine Psychologie* is worthy of a warm welcome in this country. It is quite impossible to read this book and to remain unimpressed by the remarkable parallelism between Stern's systematic outlook and the central tradition of British psychology. Stern's *Personalistik*, in its psychological aspects, is essentially the position of Ward restated in the context of contemporary experimental science. But this psychology, unlike Ward's, is no 'generalized' psychology. Stern's early training in Applied Psychology has left him in no doubt as to the supreme reality and significance of individual differences. The roots of 'personalistics' clearly go back to the classical researches embodied in the *Differentielle Psychologie* and the celebrated series of *Beiträge zur Psychologie der Aussage*. And these, in their turn, may surely be traced to the epoch-making contribution of Galton.

William Stern's theory of memory, elaborated at great length in Part III of this book, is of particular interest to English readers. In the first place, he rejects out of hand the outworn doctrine of memory traces, and introduces in its stead a theory of psychical dispositions almost identical with that of Stout. In the second place, he accepts as fundamental the distinction between 'bound' and 'free' ideas—a distinction which most of us have come to associate so intimately with the writings of Ward, Stout, McDougall and Hobhouse. And in the third place, his treatment of remembering itself, with its emphasis upon fluidity and construction, challenges serious comparison with the more recent writings of Bartlett. The insufficiency of associationism, moreover, was clearly appreciated by Stern, and the necessity for a dynamic approach to the central problems of psychology is exploited with a persuasiveness reminiscent of Stout and a vigour unmatched in the most outspoken polemics of McDougall. William Stern is surely entitled to a high place in a generation of psychologists whose principal contribution has lain in freeing their science from the arid sterility of the Leipzig tradition.

Dr Spoerl deserves our warmest congratulations upon the competence and skill with which he has carried out a difficult piece of translation. The text of the original German edition (which appeared in 1935) has been considerably revised, the author tells us in his Preface, "to meet the requirements of English and American readers". The bibliography, too, has been extensively modified, presumably with the same somewhat nebulous end in view. This appears in six parts, and is comprehensive enough, though much of the material quoted has little explicit reference to the text which precedes it.

It is to be hoped that the publishers will see their way to preparing an English edition of this book, more dignified in appearance and less costly in price. Such a step would undoubtedly assist in realizing the author's wish that "the English form of the book may strike a sympathetic note".

Human Understanding and its World. By K. W. MONSARRAT. University Press of Liverpool (London: Hodder and Stoughton). 1937. Pp. xx+480. 15s. net.

Human Powers and their Relations. By K. W. MONSARRAT. University Press of Liverpool (London: Hodder and Stoughton). 1938. Pp. xvi+289. 10s. 6d. net.

These two books have to be considered together, and represent a work planned and carried out on the grand scale. Together they range over the greater part of human history and they cover a vast number of specialized problems in both physical and biological science. In a very wide sense indeed they may be considered as a contribution to psychology, for they deal with human ideas and with the relation of these to the social and material world. But the empirical psychologist looking for indications as to exactly how the ideas which are discussed at length are formed, or even precisely what they are to be taken to mean, will be somewhat disappointed.

The sub-title of the first volume is *A Study of Societies*. There are three parts named respectively *Illustration of Imagination in History*, *The Scientific Image*, and *The Making of the Image and its Significance*. Image and imagination here are general words for any process of contact with and interpretation of the external world. The general contention is that any human individual knows himself and his fellows in two ways: directly, by way of action, and indirectly through the medium of some sensorial equipment. From this he tends to look upon himself and the world as constituted by a duality of 'power' and 'matter'. But this is a mistake, because the second interpretation is really derivable from the first, and as soon as the relativity of all indirect ways of 'knowing' is recognized, it becomes clear that all nature must be regarded as 'purely dynamic'. Not only so, but, apparently, we have to admit that all world 'assemblies' are made up of 'power units' of the same order though of very varying behaviour characteristics. That units of this kind should join into dynamic systems is a 'common terrestrial event', and wherever such systems are found it is also found that apparent behaviour is the expression of a 'universal inclination to states of balance' and a tendency to seek more and more stable and wide integrations.

The second volume is occupied by a prolonged discussion of the ways in which the author believes these hypotheses to be achieved by man through the exercise of his 'Knowing', and of their applications to fairly simple physical and biological problems, and to immeasurably complex problems of human social organization.

It is impossible not to admire the pertinacity and skill with which the author has carried through an immense task. Hardly any readers will be able to follow him all the way because the specialized knowledge demanded is too great, but most intelligent readers will find much to interest and stimulate them. Dr Monsarrat is probably much too uncritical of views which appear to him to fit in well with his general picture, and he is bound to slur over many of the great difficulties. He does not define his terms very clearly, and, as has been said already, the actual psychology of idea-formation is not really discussed at all. Nevertheless he has accomplished an enormous task, and the effort to relate and read some general meaning into large tracts of data which commonly are considered only in a specialist manner was well worth while.

The American Criminal. By ERNEST ALBERT HOOTON, with the collaboration of the Statistical Laboratory of the Division of Anthropology, Harvard University. Vol. 1: *The Native White Criminal of Native Parentage*. Harvard University Press. 1939. Pp. 309, with an appendix of tables. \$10.00.

"This", say the publishers, "is the first of three large monographs which will present fully the materials and results of a survey of race and nationality in relation to crime in the United States, carried out by the division of anthropology of Harvard

University with the cooperation of the Massachusetts State Department of Mental Diseases." The work is claimed as "the most extensive anthropological study of criminals ever made", and it certainly appears that the claim is not exaggerated. Tremendous in size as the work is—a great amount of space is occupied by the very complete presentation of statistical tables—it is throughout very clearly and cleverly written and is full of interest. There is a good historical introduction in which Lombroso receives a certain amount of well merited rehabilitation, and Goring's magnificent work is treated with discriminating sympathy. The general bulk of the discussions concern sociological and morphological characteristics, with considerable reference also to topographical, occupational and racial factors. Everything is admirably summarized at the end. The English price of the work is 42s. (Oxford University Press), and it seems highly desirable that a somewhat cheaper edition should be published, in which only a selection of the statistical material need be given. The full monograph contains 400 or more pages of tables.

Crime and the Man. By ERNEST ALBERT HOOTON. Harvard University Press (London: Oxford University Press). 1939. Pp. ix + 403. 16s. net.

This is not a summary of the same author's *American Criminal*. It is an account for the general reader of the whole investigation of which that work presents only one-third. It is free from tables, but contains a lot of highly ingenious drawings. If any reader extracts more from these drawings than is meant it will be his own fault, for precisely what is meant is clearly explained in the Preface. The work deals with crime and criminals in North America. It is informed with fact and written with the most admirable good humour and incisiveness. It is rather sociological and morphological than psychological, as psychology in this field has usually been understood. But it is most highly desirable that it should receive serious study from psychologists. It does make proposals and make them incisively. And it does present an aspect of the problems involved which people who are wholly preoccupied with individual studies very often and somewhat disastrously overlook.

Personal Aggressiveness and War. By E. F. M. DURBIN and JOHN BOWLBY. London: Kegan Paul, Trench, Trubner and Co., Ltd. 1939. Pp. viii + 154. 5s. net.

This book contains an essay which appeared originally in 1938 as part of a symposium entitled *War and Democracy*, edited by E. F. M. Durbin and G. E. G. Catlin. The argument of the main part of the essay is that the inherent aggressiveness of the individual is thwarted and repressed when it runs contrary to the dictates of society. The pain of this repressed aggressiveness and the feelings of guilt which arise from it are relieved by displacement and projection upon some real or imagined enemy of the social group, the scape-goat; and the tendency is enhanced by the 'animistic' tendency to attribute disasters and misfortunes to the antagonism and evil machinations of certain persons—among primitive tribes of evil spirits, and in modern civilized societies of Jews, capitalists, etc. This thesis is most ably exposed, and is well supported by the evidence adduced in the appendix, which has been drawn from the psychological study of children, from psycho-therapy, from anthropological studies of primitive societies, etc. Unfortunately, though the diagnosis of the disease seems convincing, the remedies suggested—tentatively, it is admitted—are less so. When the authors pass from observation to prediction, and attempt to assess the relative values of pacifism *v.* collective security, and the effectiveness of various methods of educational treatment, they seem to lose their objectivity and psychological insight. Social psychology is not yet in a state sufficiently advanced to allow of reliable predictions as to the behaviour of the complex groups of modern civilized society.

Causes of Crime: Biological Theories in the United States, 1800-1915.

By ARTHUR E. FINK. Philadelphia: University of Pennsylvania Press (London: Humphrey Milford). 1938. Pp. xi+309. 14s. net.

This is an interesting and well-documented study of theories about the causation of crime current in the U.S.A. throughout the period under review. It shows a general development from dogmatic and general theorizing to a more scientific and, in many ways, a more individualistic type of approach. The author shows also that towards the end of the period the individual studies are complemented by the very widespread use of statistical techniques. The general upshot of the whole survey is that the criminal is now to be regarded neither wholly as a product of environmental forces nor as a result of inborn biological factors. He is not to be characterized simply by making a list of composite traits, but must be considered, like any other person, as a 'functioning, integrated' personality. A very excellent bibliography is included.

Taste and Temperament. By JOAN EVANS. London: Jonathan Cape, Ltd.

1939. Pp. 128+48 illustrations. 10s. 6d. net.

The author of this book advances an interesting and suggestive theory, namely that artistic taste is a function of individual temperament and that each temperamental type creates and appreciates its characteristic form of visual art. But persuasively as it is put forward, the thesis seems to be entirely speculative. The particular typology selected—that of quick and slow introvert and extrovert—is not based upon any kind of enquiry or investigation into individual temperament or reaction to art forms, but upon a rather superficial study of the typologies previously suggested by psychologists and others. The analysis of the various art forms and of their manner of appeal is masterly; and the evidence as to the particular form most characteristic of the artists of a given temperamental type is fairly convincing. But it does not seem to follow that there is the same correspondence of taste and temperament in artistic appreciation as in artistic creation. Indeed, the Jungian theory might lead us to suspect the opposite. But even though the author's argument is largely speculative, it does lead to the valuable conclusion that it is useless to try to persuade people, and especially children, to appreciate forms of art to which they are temperamentally antipathetic. Unfortunately the other suggestions cannot be followed up until more is known of temperamental types and of their characteristic tastes.

The book is beautifully printed and illustrated.

Practical Child Psychotherapy. By CURT BOENHEIM. London: John

Bale Medical Publications, Ltd. 1938. Pp. 177. 10s. 6d. net.

Starting from the standpoint that "the body of the child is largely the centre around which conflicts and their resulting neurotic reactions appear to be grouped", the author describes the diagnosis and treatment of the various somatic disorders in children which are mainly or partly psychological in origin. He stresses the importance of observing and where possible modifying environmental difficulties; and also of allaying the actual somatic symptoms with drugs, rest, exercise, etc. Various methods of psychological treatment are also advised and described, but rather sketchily. In fact, as its title implies, the book will probably be of more value to the medical practitioner than to the lay psychologist or psycho-therapist. But its usefulness is somewhat impaired by the stilted and involved language of the translation. A good bibliography is appended, though it is suggested that more English translations of German books might be included.

Modifications proposed for British use of the Revised Stanford-Binet Tests of Intelligence, in Measuring Intelligence, by Lewis M. Terman and Maud A. Merrill. By THE SCOTTISH COUNCIL FOR RESEARCH IN EDUCATION, 46, Moray Place, Edinburgh, 3. 1939.

The Scottish Council for Research in Education have prepared duplicated copies (Copyright, All Rights Reserved) of their proposed modifications in the procedure and scoring of the new Revised Stanford-Binet Tests of Intelligence (Terman and Merrill), designed to render them more suitable for British children. Copies are available to accredited testers on application to the Research Council. These modifications have the approval of Prof. Terman and Dr Merrill. They agree with some of the modifications suggested by the English Committee under Prof. Burt, but do not go as far as these. The original standardization is retained.

The Nineteen Thirty Eight Mental Measurements Year Book of the School of Education, Rutgers University. Edited by O. K. BUROS. New Brunswick, Jersey, U.S.A.: Rutgers University Press. 1938. Pp. xv+415. \$ 3.00.

This book, which is published to some extent as an experiment, gives a very complete and comprehensive list of descriptive references to current tests and their uses. It deals also with Mental Measurement books, with research and statistical methodology books, with regional programme reports about testing technique and practice, and gives both a periodical directory and index and a publishers' directory and index. There are also full indexes of all titles dealt with and of all names of authors referred to. But the book is more than just a list. All the articles and books dealt with are briefly described, and in general the descriptions are very clear and good. Further, in the cases in which articles and books have been reviewed, a selection of the reviews is abstracted and their relevant comments indicated. Consequently it is possible to get a fairly good idea both of the matter of the publications listed and of its value as a contribution to psychological theory and practice. The amount of work involved, both on the part of the editor, and on the part of his many contributors, must have been very great indeed. It is to be hoped that the book will find a place in the libraries of all psychological departments and research centres, and that the editor will be encouraged to continue a task which, though it must be arduous, will be of undeniable service to all students who wish to engage in research in the general field covered.

Foundations of Educational Psychology. By PETER SANDIFORD. London: Longmans, Green and Co., Ltd. 1938. Pp. xv+464. 12s. 6d. net.

The foundations of educational psychology seem to be mainly biological and physiological, with a superstructure firmly built up from biometry, psychometry and Watsonian behaviourism. A rather airy top floor is constituted by the nature of intelligence and personality, but the Freudians and the Gestalt psychologists have been driven into the attics, and the child study of Bühler, Isaacs, etc., seems to have been thrown out of the window. The specialist would find himself at home in the lower part of the building, and indeed might employ himself usefully there (the study of the inheritance of mental characteristics is a valuable one). But the upper part is more suitable for the tyro. This method of construction may be instructive to the experimental psychologist; is it really helpful to the student of education?

Manhood in the Making. Edited by T. F. COADE. London: Peter Davies. 1939. Pp. xi+347. 10s. 6d. net.

There has been recently a rapid spate of books consisting of heterogeneous collections of essays or articles by various authors, with or without some attempt at editing into coherent wholes. The editor of this latest example seems to have endeavoured to select his authors in accordance with a pre-arranged scheme, so that they might demonstrate exhaustively the various aspects of the boy's development, in relation to his environment and his education. Unfortunately no connected system of ideas is ascertainable running through the various essays; and contradictions are frequent and unexplained. Moreover, although many of the essays, taken singly, are interesting, their cumulative effect is wearisome; for *ex cathedra* pronouncements are as a rule unsupported by evidence, and vague generalizations yield few concrete and specific examples. There are of course certain honourable exceptions to these strictures, of whom Prof. Burt affords the most striking. But we are forced to conclude that this type of book is of doubtful value unless there is some real collaboration, or some fundamental unity of ideas, between the different authors.

The Rediscovery of Man. By HENRY C. LINK. London: Macmillan and Co., Ltd. 1939. Pp. ix+257. 7s. 6d. net.

Describes how young men and women, by playing competitive games, can become good mixers, and hence Make Good.

A Study of Jealousy as differentiated from Envy. By T. M. ANKLES. London: Arthur H. Stockwell, Ltd. 1939. Pp. 112. 3s. 6d. net.

This book contains the account of a piece of research work carried out for the Academic Postgraduate Diploma in Psychology. The material was obtained in the main by asking casual acquaintances to say what they 'felt' by jealousy, and encouraging them to introspect on this subject; in every case personal experiences were described, although they were not asked for directly. This material was then interpreted by the author, on a more or less Freudian basis. Both procedure and conclusions are interesting and suggestive, though possibly too superficial for so difficult a subject. Unfortunately the author's style of writing is so incoherent and lacking in lucidity that it is impossible to evaluate his work with any justice.

The Industrial Worker: A Statistical Study of Human Relations in a Group of Manual Workers. By T. N. WHITEHEAD. London: Oxford University Press. 1938. Vol. I, pp. 265; Vol. II, 81 diagrams with statistical explanations.

In this book Prof. Whitehead has carefully analysed the voluminous data obtained from five years' observation of a special group of repetition workers. He has already given less full accounts of this experiment in shorter papers, but the present book is definitely for the expert who wants fully to understand his methods and judge the value of his results. It is impossible in a short notice to give an adequate account of such a detailed piece of work. Briefly the thesis put forward is that statistically determined social relations modify activity as measured by industrial output. He shows that, within the limits of adaptation, physical changes in environment play an insignificant part in working efficiency compared with the part played by social relationships.

The book is interesting because it breaks new ground in the methods of social psychology, bringing it definitely into line with the statistical methods adopted by experimental psychology. At the same time it advances beyond the stage of enumerating statistical relationships by relating them to the constantly changing social relations of the members of an experimental group.